DOCUMENT RESUME

ED 391 689 SE 057 767

AUTHOR	Posamentier, Alfred S.; Schulz, Wolfgang
TITLE	The Art of Problem Solving: A Resource for the Mathematics Teacher.
REPORT NO	ISBN-0-8039-6362-9
PUB DATE	96
NOTE	473p.
AVAILABLE FROM	Corwin Press, Inc., 2455 Teller Road, Thousand Oaks, CA 91320.
PUB TYPE	Books (010) -- Collected Works - General (020) -- Guides - Classroom Use - Teaching Guides (For Teacher) (052)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Cooperative Learning; *Educational Resources; Elementary Secondary Education; *Mathematical Concepts; Mathematics Teachers; *Problem Solving

ABSTRACT

This book is designed to give mathematics teachers a host of interesting and useful ideas thereby raising their consciousness level and enabling an enrichment of the mathematics instruction program. The chapters in this book capture a broad spectrum of ideas in the area of mathematics problem solving. Chapters are: (1) "Strategies for Problem Exploration" (Ewen, Ira); (2) "Unconventional Problem-Solving Strategies in Mathematics Instruction" (Posamentier, Alfred); (3) "Interest Grabbers: Exciting Motivational Problems with Punch and Personality" (Conrad, Steven); (4) "Check the Answer, Please!" (Salvadori, Mario); (5) "The Logic of Error" (Akin, Ethan); (6) "Trial and Success" (Paul, Fred); (7) "Reduce, Expand, and Look for a Pattern" (Krulik, Stephen; and Rudnick, Jesse); (8) "The Pigeonhole Principle for Problem Solving" (Posamentier, Alfred; and Lee, Wei); (9) "Handling, Seeing, and Thinking Experiences in Mathematics" (Maletsky, Evan); (10) "Problem Solving as a Continuous Principle for Teaching: Suggestions and Examples" (Humenberger, Hans; and Reichel, Hans-Christian); (11) "Another View of Combinatorics (or Counting Without Really Counting)" (Moresh, Stephen); (12) "Problem Solving by the Use of Functions" (Schulz, Wolfgang); (13) "Symmetry Saves the Solution" (Singmaster, David); (14) "An Application of Congruence Transformations in Problem Solving" (Trojak, Jan); (15) "Graph Theory: Tools to Solve Mathematical Problems" (Lee, Wei); (16) "A Different Solution for Problems With Extreme Values: A Didactic Journey Into the World of Jakob Steiner's Ideas" (Kiesswetter, Karl; Stowasser, Roland; and Haapasalo, Lenni); (17) "The Problem of the Duplication of a Cube" (Kaiser, Hans); (18) "Solving Mathematical Problems Using Game Strategies" (Kauke, Marion; and Ziller, Sabine); (19) "Cooperative Learning Approaches to Mathematical Problem Solving" (Hartman, Hope); and (20) "Problem Solving and the Mathematically Gifted Student: A Psychological Perspective" (Rollett, Brigitte). Appendices include sources for problems and readings on problem solving. (JRH)

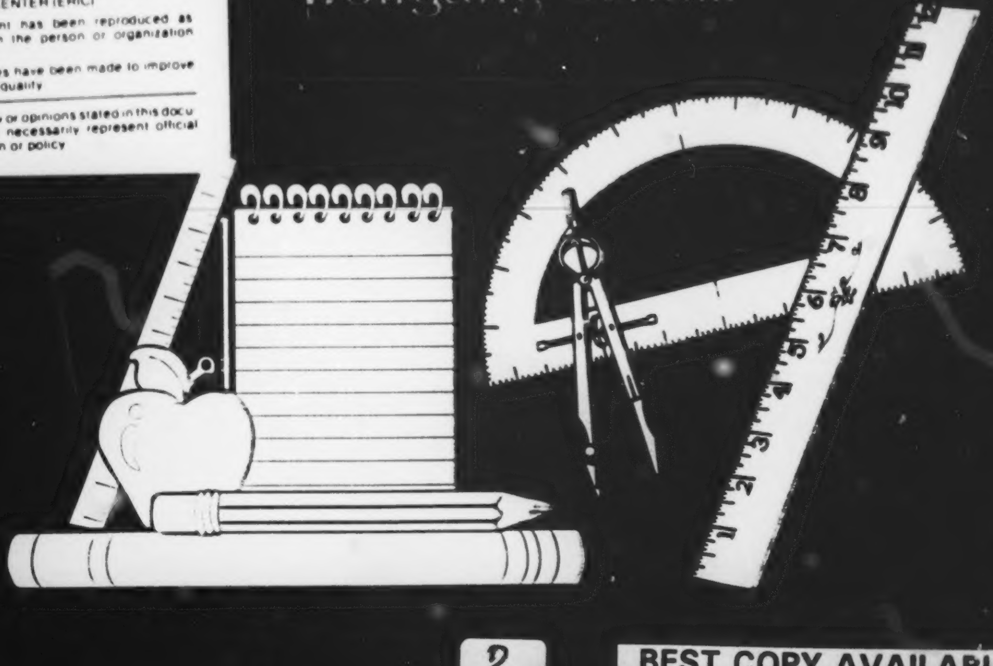
ED 391 689

"PERMISSION TO REPRODUCE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY
G. Alkema
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Art of Problem Solving

Editor
Alfred S. Posamentier
Associate Editor
Wolfgang Schulz

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy



BEST COPY AVAILABLE

The Art of Problem Solving

To our children,
Lisa Joan, David Richard,
Birgit, and Beate,
who provide us with constant reinforcement
to hone our problem-solving skills

The Art of Problem Solving

A Resource for the Mathematics Teacher

Editor
Alfred S. Posamentier
Associate Editor
Wolfgang Schulz

CORWIN PRESS, INC.
A Sage Publications Company
Thousand Oaks, California

Copyright © 1996 by Corwin Press, Inc.

All rights reserved. Use of the handouts, resources, and sample documents in each chapter is authorized for local schools and noncommercial entities only. No other part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

For information address:

Corwin Press, Inc.
A Sage Publications Company
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications Ltd.
6 Bonhill Street
London EC2A 4PU
United Kingdom

SAGE Publications India Pvt. Ltd.
M-32 Market
Greater Kailash I
New Delhi 110 048 India

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

The art of problem solving: A resource for the mathematics teacher/
editor(s), Alfred S. Posamentier and Wolfgang Schulz.
p. cm.
Includes bibliographical references.
ISBN 0-8039-6361-0 (cloth: acid-free paper) —
ISBN 0-8039-6362-9 (pbk: acid-free paper)
1. Problem solving. I. Posamentier, Alfred S. II. Schulz,
Wolfgang, 1943 Feb. 3-
QA63.A78 1996
510'.76—dc20 95-32522

This book is printed on acid-free paper.

95 96 97 98 99 10 9 8 7 6 5 4 3 2 1

Corwin Press Production Editor: Gillian Dickens
Corwin Press Typesetter: Janelle LeMaster

Contents

Introduction vii

1. Strategies for Problem Exploration 1
Ira Ewen

2. Unconventional Problem-Solving Strategies in Mathematics Instruction 83
Alfred S. Posamentier

3. Interest Grabbers: Exciting Motivational Problems With Punch and Personality 101
Steven R. Conrad

4. Check the Answer, Please! 121
Mario Salvadori

5. The Logic of Error 131
Ethan Akin

6. Trial and Success 141
Fred Paul

7. Reduce, Expand, and Look for a Pattern 149
Stephen Krulik and *Jesse A. Rudnick*

8. The Pigeonhole Principle for Problem Solving 159
Alfred S. Posamentier and *Wei Lee*

9. Handling, Seeing, and Thinking Experiences in Mathematics 173
Evan M. Maletsky

10. Problem Solving as a Continuous Principle for Teaching: Suggestions and Examples 199
Hans Humenberger and *Hans-Christian Reichel*

11. Another View of Combinatorics (or Counting Without Really Counting) 233
Stephen E. Moresh

12. Problem Solving by the Use of Functions 257
Wolfgang Schulz

13. Symmetry Saves the Solution 273
David Singmaster

14. An Application of Congruence Transformations in Problem Solving 287
Jan Troják

15. Graph Theory: Tools to Solve Mathematical Problems 313
Wei Lee

16. A Different Solution for Problems With Extreme Values: A Didactic Journey Into the World of Jakob Steiner's Ideas 349
Karl Kiesswetter, Roland J. K. Stowasser, and *Lenni I. Haapasalo*

17. The Problem of the Duplication of a Cube 371
Hans K. Kaiser

18. Solving Mathematical Problems Using Game Strategies 383
Marion Kauke and Sabine Ziller

19. Cooperative Learning Approaches to Mathematical Problem Solving 401
Hope J. Hartman

20. Problem Solving and the Mathematically Gifted Student: A Psychological Perspective 431
Brigitte A. Rollett

Sources for Problems 445

Readings on Program Solving 453

About the Editors 457

About the Contributors 459

Introduction

From the earliest times, a cornerstone of mathematics has been problem solving. Too often, solving a problem is seen as a means to an end rather than as an end in itself. In the collection of ideas presented in this book, we hope to change that perception. Here, problem solving is the main theme. It is presented from the very practical to the theoretical, from the common to the glitzy, from the purely mathematical (viewing the beauty of mathematics for its own sake) to the pedagogical and psychological considerations surrounding the "sharing" of this important discipline.

This book is intended for mathematics teachers. It was conceived as an effort to bring them a host of interesting and useful ideas, thereby raising their consciousness level and enabling an enrichment of the mathematics instruction program.

To ensure a wide variety of ideas without repeating notions commonly found in current literature, we drew from a team of authors representing a wide array of geographical areas as well as a broad experiential range. The contributing authors in this book are professionals who, in a variety of ways, influence mathematics education internationally and range in age from the mid-30s to the upper 80s.

A key stipulation for the authors, who were given free rein in terms of space and topics, was to discuss their favorite ideas in the realm of problem solving in the most enticing manner possible, so as to make the book enjoyable and entertaining as well as useful. This inevitably has led to an occasional example being used more than once. We chose not to change these duplications because their replacement or removal would have altered the individual contribution. Authors were urged to write in an informal and easily understood style.

The book can be used in a number of ways. For the most part, the chapters are independent of one another and, therefore, can be read in any order. Our order of presentation was chosen to provide a meaningful guide through the extensive and complex world of problem solving. Naturally, other orderings not only are possible but may even serve some readers better, so readers should feel free to personalize the order in which they read the chapters. To facilitate other reading possibilities we offer the reader two marginal guides: One will indicate the problem-solving strategies being used, and the other will allow an easy subject matter identification.

Throughout the book, a variety of problem-solving techniques are introduced, discussed, and illustrated. Similar techniques can manifest themselves somewhat differently within various contexts. A reader might find it helpful, after being stimulated by a particular technique, to locate other applications of those particular problem-solving methods by means of the circled numerals in the margins. The numbers and their corresponding techniques are as follows:

1. Working backwards
2. Finding a pattern
3. Adopting a different point of view
4. Solving a simpler analogous problem
5. Considering extreme cases
6. Visual representation (diagram, table, chart, etc.)
7. Intelligent guessing and testing
8. Determining necessary and sufficient conditions
9. Sequencing
10. Specification without loss of generality
11. Systematically accounting for all possibilities
12. Using a computer
13. Deductive reasoning
14. Organizing data
15. Approximating
16. Determining characteristics of objects
17. Specializing
18. Generalizing

The boxed "signposts" in the margins allow the reader to search for problem-solving discussions by subject matter strategy. The categories were selected to reflect broadly the mathematics taught in schools today. This may not always be the best way to consider the topic of problem solving in mathematics, but because we would like to make this book as reader friendly as possible, we are providing these signposts as an additional pathfinder for those readers bent on pursuing a subject-matter focus. They should also prove useful for those unconvinced readers

ALG	Algebra
ARITH	Arithmetic
COMB	Combinatorics
GAME	Games
GEN	General application
GEOM	Geometry
LOGIC	Logic
NUMB	Number theory
PROB	Probability
STAT	Statistics
TOPOL	Topology
TRIG	Trigonometry

Figure 1

who, despite our efforts to highlight problem solving as a central theme, still want to investigate mathematics by topic, or for those readers who are looking for clever applications of problem solving for a particular topic in order to enhance their instruction. The key to these signposts follows (see Figure 1).

The chapters capture a broad spectrum of ideas in the area of mathematics problem solving. We begin with "Strategies for Problem Exploration," a truly engaging "conversation" with the reader on how to recognize problems and how to consider appropriate solution strategies. This chapter serves as a good introduction to the book, since it covers most of the strategies listed above. This comprehensive chapter is followed by "Unconventional Problem-Solving Strategies in Mathematics Instruction," a chapter highlighting problem-solving strategies that exhibit surprising and often elegant solutions, particularly where they may not necessarily be expected. Through demonstration of their use in a somewhat ideal application, it is intended that these problem-solving methods will become part of the reader's arsenal of problem-solving techniques. To further exhibit the motivational value of clever problems coupled with even cleverer solutions, the next chapter, "Interest Grabbers: Exciting Motivational Problems with Punch and Personality," offers a plethora of problems from many areas of school mathematics. This chapter will be truly entertaining to read, and, as the author says, the reader should be an active participant, not a "spectator."

The succeeding chapter, "Check the Answer, Please!" written by a world-renowned engineer, discusses how real-life problem solutions that are not carefully inspected can lead to embarrassing situations. It also demonstrates that mathematical problems exist which have very unexpected, or "impossible," answers that are, in fact, correct. The theme is that there is more to checking an answer than its reasonableness!

The three chapters that follow consider the problem-solving process from a more pedagogical point of view. "The Logic of Error" considers some frequent types of mathematical mistakes which can considerably frustrate problem-solving efforts, while "Trial and Success" and "Reduce, Expand, and Look for a Pattern" focus on the instructional process desirable for strengthening problem-solving skills.

The next two chapters, "The Pigeonhole Principle for Problem Solving" and "Handling, Seeing, and Thinking Experiences in Mathematics," present specific and somewhat novel approaches to attacking mathematics problems which harmlessly seem to evade simple solutions (or at least methodical approaches). The tools developed here exhibit their potential power by their application to a variety of mathematical situations that do not necessarily call for these methods.

The chapter "Problem Solving as a Continuous Principle for Teaching: Suggestions and Examples" supports the notion that problem solving is not a theme reserved only for gifted students. Rather, through appropriately selected problems, less gifted students can also reap substantial benefit from such endeavors. The authors of this chapter stress the importance of investigating the sufficiency of information given about a problem, with the belief that this sort of problem analysis develops a deeper insight into the problem-solving process.

"Another View of Combinatorics (or Counting Without Really Counting)" presents a delightful, informal approach to problem solving in the realm of probability and counting. A greater facility and a deeper insight into problems in this field is an aim of this chapter.

In "Problem Solving by the Use of Functions," functions are used to demonstrate some special problem-solving techniques. The author uses a variety of examples to make the point that where functions can be used to solve a particular problem, some specific solution methods are available.

The author of "Symmetry Saves the Solution," well known in the realm of recreational mathematics, poses several problems of a recreational nature and uses them as a vehicle for demonstrating some elegant solution methods. The chapter's use of unexpected methods is intended to prompt readers to develop new problem-solving procedures.

"An Application of Congruence Transformations in Problem Solving," as the title suggests, uses transformation techniques to solve geometric problems. The material here begins with a good knowledge of the high school geometry course.

"Graph Theory: Tools to Solve Mathematical Problems" is an attempt to demonstrate how this relatively modern branch of mathematics can be used to solve problems typically attempted through other means, for such topics as combinatorics, games, and number theory. One of the advantages of using graphs as a problem-solving tool is that they allow for visualization of ideas which otherwise remain very abstract.

The next two chapters refer to the history of mathematics. In "A Different Solution for Problems With Extreme Values: A Didactic Journey Into the World of Jakob Steiner's Ideas," the concept of using extreme values to solve equations is demonstrated primarily in a geometric context. A historical perspective is presented in "The Problem of the Duplication of a Cube," where specific problem-solving experiences from the past are presented and explored.

The remaining chapters—"Solving Mathematical Problems Using Game Strategies," "Cooperative Learning Approaches to Mathematical Problem Solving," and "Problem Solving and the Mathematically Gifted Student: A Psychological Perspective"—discuss problem solving from a psychological point of view, considering not only the psychological aspects but also the pedagogical ones. The presentations allow us to continue to learn from experience.

Our two primary goals are to win over the mathematics teacher to believe in the notion that the role of problem solving in the classroom is the key component in the instructional program, and to provide sufficient applications to motivate the teacher to experiment in the classroom.

—*Alfred S. Posamentier*
New York

—*Wolfgang Schulz*
Berlin

1

Strategies for Problem Exploration

IRA EWEN

Many of the most significant problem situations we encounter in life are ill-defined. Often it is not immediately clear which questions should be posed and in which order they should be addressed; moreover, the nature of the desired outcomes may be clouded. Problem situations require thorough study by someone skilled in coping with murky and unfamiliar situations. The process of studying ill-defined problem situations—posing and prioritizing questions, investigating and modifying courses of analysis, investigating and developing alternative solutions, generalizing results, and going off in new directions—would benefit from a rubric more suggestive than "problem solving." Our name for this process is *problem exploration.*

In this extended chapter, we will identify and illustrate 13 major problem-exploration strategies and attempt to shed light on how they may be used in structured and ill-defined contexts:

1. Working Backwards
2. Finding a Pattern
3. Adopting a Different Point of View
4. Solving a Simpler Analogous Problem

5. Considering Extreme Cases
6. Visual Representation
7. Intelligent Guessing and Testing
8. Determining Necessary or Sufficient Conditions
9. Sequencing
10. Specification Without Loss of Generality
11. Systematically Accounting for All Possibilities
12. Using a Computer
13. Deductive Reasoning

①

1. Working Backwards

1.1. Definition and Discussion

The phrase "working backwards" has a pejorative implication for many people who are mathematically ignorant. For them, it implies somebody not knowing where they are going and wandering in a wayward direction; nevertheless, working backwards is among the most significant of all procedures for exploring ill-structured problem situations. We are working backwards when we begin the analysis of the problem not from the given information or a seemingly natural starting point but from a desired final result. We then attempt to find a more easily accessible situation that will yield that final result and then a second easily accessible situation that will yield the result at the first stage of our backward progress. This procedure is continued until we reach a condition that is known, given, or immediately accessible. When would we work backwards? The key decision is whether we have a single natural starting point, as contrasted to a single natural final result. If there are many possible final results that can solve the problem that we've defined, and a small number of natural starting points, then we should work forward in a natural and familiar way. When there is a single final result and multiplicity of starting points, it is almost always better to work backwards. Let us clarify the way we work backwards through a number of examples.

1.2. Planning an Automobile Route From New York City to Urbana, Illinois

NUMB
GEOM

Imagine you are an employee of the American Automobile Association whose job it is to create automobile travel maps for members. A professor from City University calls to request an automobile map for a trip from New York City to the University of Illinois in Urbana. Let us follow your thinking. You note that there are five natural automobile routes for leaving New York City on a trip west: the Verrazzano Bridge, the Holland Tunnel, the Lincoln Tunnel, the George Washington Bridge, and the Tappan Zee Bridge. From each of those departure points, there are many routes west. Where should you start? Your experience tells you that you should start at none of them. Since there are five routes from New York City, each leading to a network of highways, but a single small destination point, the problem cries out for working backward. To reach Urbana, we must have been traveling west, so let's plot the return trip. We leave Urbana and travel to the first major east-west highway. We take it eastward, seeking major interstate highways that are also east-west, and continue working our way east, changing highways as necessary, toward the New York City megalopolis. No matter how you do this, if you are careful, you are led back through one of the five access routes to New York City. You send the professor a map indicating the route you have drawn, and you have simply and clearly found an acceptable automobile route from New York City to Urbana. Note how natural this is. There was nothing wayward or blundering about the procedure. Working backward was the natural way to go.

1.3. Games of "27-Depletion (1, 2, 3, 4)"

② GAME ⑪ ⑬

You are challenged to play a set of depletion games with the following common rules:

1. There are exactly two players.
2. Turns alternate.
3. At each turn, a player removes 1, 2, 3, or 4 counters from a pile that was initially 27 counters.
4. The game ends when all counters have been removed.

There are several natural fifth rules specifying the objective of the game. Let us examine them in turn.

Rule 5A: The player who takes the last counter wins.

Note that it is not obvious whether it is better to go first or to go second. If you were told to go first, it would be unclear how many counters to take. If you were told to go second, it would be unclear how to capitalize on an opponent's error or how to maintain a winning position if you, as the second player, had one. We know from game theory that, because this is a finite game for two players, which cannot end in a draw, one of the two players has a perfect strategy that will lead to victory. How can we find that strategy efficiently? Note the characteristics of this problem. There are four possible first moves (taking 1, taking 2, taking 3, taking 4), but only one winning final move, namely, taking all of the remaining counters when there are 4 or fewer left. Again, this is a problem asking you to work backward. Let us follow along as we trace our backward path. To guarantee my ability to leave 0 counters at the end of my previous turn, I must leave a number of counters that has two characteristics: (a) My opponent cannot take all of them, yet (b) whether he takes 1, 2, 3, or 4, I can take all of remaining counters. A moment's reflection will indicate that the number I must leave on my next-to-last turn is 5. If my opponent takes 1, I take 4; if my opponent takes 2, I take 3; if my opponent takes 3, I take 2; and if my opponent takes 4, I take 1. Note that, for example, leaving 6 on my next-to-last turn would fail, because my opponent could take one, leaving me in a dilemma with no resolution. We have taken our first backward step.

To guarantee my ability to leave 5 counters in my next-to-last turn, I ask what number of counters I must leave on my turn before that to guarantee my access to leaving 5 counters on my next-to-last turn and deny that access to my opponent. Brief reflection will show that I must leave 10 counters on my next-to-last turn, for reasons analogous to those described above to explain why leaving 5 guarantees a win. A simple pattern has emerged. On the backward route, I must leave 0 counters, 5 counters, 10 counters, 15 counters, 20 counters, and 25 counters.

All is clear. To win the game, I must go first and take 2 counters. At each subsequent turn, I must leave a number of counters that is a multiple of 5. By doing this, I guarantee that after my last turn, I will have left none and will have won the game.

Rule 5B: The player who takes the last counter loses.

Once again, we are confronted by a situation with four possible opening moves but only one possible final move. I win the game by leaving 1 counter for my opponent. Arguments analogous to those explained in the discussion of Rule 5A indicate that on my backward analysis of this game I must leave 1 counter, 6 counters, 11 counters, 16 counters, 21 counters, and 26 counters. Again all is clear. I win by going first and taking 1 counter. In subsequent turns, I always leave a number of counters that is exactly one more than a multiple of five. On my final turn, I will leave 1 counter and will have won the game.

Rule 5C: The player having an even number of counters
when all the counters have been taken wins.

This time it is neither clear how to start nor how to end. If I leave 0 counters and have an even number, I will win; but I will also win when I have an even number of counters if I leave one. A productive backward analysis would continue with the question, "What is the minimal number of counters I could leave to guarantee a win if I am holding an odd number of counters?" Reflection will show that the answer to that question is 5. Let us see why. If I hold an odd number of counters and leave 5, my opponent has only losing choices: If he takes 1, I take 3; if he takes 2, I take 3; if he takes 3, I take 1; and if he takes 4, I take 1. In each case, I reverse the parity of my holding from odd to even and leave him with a fatal 0 or 1 counter. To work backward from this end position, it is necessary to realize that there are three parts to the winning end position: two consecutive numbers if I hold an even number of counters and a single number if I hold an odd number of counters. Where do we step on our backward path? The analysis requires significantly deeper thought than the backward steps in the previous two games. If I am to hold an even number of counters, I expect that my first backward step will lead me to two consecutive numbers. What are they?

Analysis reveals them to be 6 or 7. Let us see why. Suppose I hold an even number of counters and leave 6. If my opponent takes 1, I take 4, maintaining the parity of my holding, and guarantee a win. If he takes 2, I take 4, winning immediately. If he takes 3, I take 2, maintaining parity and forcing him to take the last one. Finally, if he takes 4, I take the last 2 and win.

Let us now analyze what happens if I hold an even number of counters and leave 7. If my opponent takes 1, I take 1, reversing the parity of my holding from even to odd and leaving 5, which is a previously analyzed winning position. Were my opponent to take 2 from the 7, I would take 4, maintaining the parity of my holding and forcing him to take the last one. Were he to take 3, I would take 4, winning immediately. Were he to take 4, I would take 2, preserving the parity of my holding and forcing him to take the last one. I note that with an even holding there is a difference of 6 between the 0, 1 pair of winning leaves and the 6, 7 pair. I conjecture the following pattern of winning leaves.

Parity of My Own Holding	*Winning Leave*
even	0, 1, 6, 7, 12, 13, 18, 19, 24, 25
odd	5, 11, 17, 23

All is clear. I must go first and take 2. Nothing else permits access to the pattern of winning leaves.

This is a difficult analysis. Working backward made it possible, though not easy, to complete the analysis.

*Rule 5D: The player having an odd number of counters
when all the counters have been taken wins.*

The analysis is left to the reader. The reader should be aware that *none* of the games discussed in this section is fair, because the player who goes first can always force a win. If the first player errs, his opponent can now always force a win. The point of giving this game as an example of working backward is to foreshadow the next section, finding a pattern.

Games are often useful because they teach important lessons with a light touch. Surely the lessons to any prospective player of any game include:

- Love the learning process as much as the thrill of winning.
- Find as many patterns as possible within the game.

⑧
⑬

- Don't expect to win against a player who knows more about the game than you do.
- Don't bet against strangers.

TRIG
LOGIC

1.4. *Proving Complicated Trigonometric Identities*

A student is asked to prove a complicated trigonometric identity such as

$$\frac{1+\cos x}{\cos x} = \frac{\cos x \tan^2 x}{1-\cos x}.$$

The student has been correctly but deceptively instructed that no proof of an equation can begin with the equation to be proved.

This is correct. Suppose the student was asked to prove (the false) equation $2 = 3$. The student reasons correctly: If $2 = 3$, then $3 = 2$ by the symmetry of equality; if $2 = 3$ and $3 = 2$, then $5 = 5$ because when equals are added to equals the sums are equal; $5 = 5$ is known to be true. The student reasons incorrectly that since he has deduced a true equation from $3 = 2$, he has proved that $3 = 2$.

The instruction that forbade beginning with $3 = 2$ was deceptive. If the student looked for reversible consequences of $3 = 2$ (reversible in the sense that each step in the logical sequence is biconditional rather than a simple consequence of the preceding step) and if the student reached a known equation, a proof would have been achieved. Notice that the equation $5 = 5$ cannot yield either $2 = 3$ or $3 = 2$ as a consequence; the student's argument is irreversible and proves nothing useful.

When analyzing complex trigonometric identities or equations or inequations, working backward is extraordinarily useful because we know where we want to end but we don't know where to begin. The need to check reversibility at each step remains.

Let us prove the identity given above. A student might say:

Provided neither $\cos x$ nor $(1 - \cos x)$ is zero,

$$\frac{1+\cos x}{\cos x} = \frac{\cos x \tan^2 x}{1-\cos x}$$

if and only if $(1 + \cos x)(1 - \cos x) = \cos^2 x \tan^2 x$
if and only if $\sin^2 x = \cos^2 x \tan^2 x$
if and only if $\sin^2 x = \cos^2 x (\sin^2 x)/(\cos^2 x)$
if and only if (remember, $\cos x \neq 0$) $\sin^2 x = \sin^2 x$.

A valid proof of the required identity would begin with $\sin^2 x = \sin^2 x$ and conclude with the required identity whenever $\cos x \neq 0$ and $1 - \cos x \neq 0$.

This method of proof is efficient, practical, and eminently valid. Note that the biconditionality of each inference is psychologically helpful because the proof *looks like* a series of inferences, which it is *not*. It would have been valid if each line were an implication of the line below it; that is, if "if and only if" were replaced by "if." We would *not* have a proof if "if and only if" were replaced by "only if." The proof actually follows from the bottom up, not from the top down.

1.5. Concluding Remarks ⑤

We have scratched the surface of working backward. However, the sequence in which we analyzed this methodology was not pedagogically optimal. We worked backward inappropriately. To help the reader understand the problem exploration strategies, it is better to provide experiences that suggest a definition of those strategies. Far superior is the procedure that we shall use in the ensuing sections. We shall give several examples of a strategy and conclude with a definition and discussion of that strategy so that reference can be made to the examples already developed. This procedure will help most readers to efficiently construct knowledge of what the strategies to be discussed mean and how they are used.

2. Finding a Pattern

2.1. The Game "Moon Is, Sun Isn't" ⑦

GAME

Two children, ages 9 and 7, are traveling in a car with their parents. The father says, "Moon is, sun isn't." The three passengers are accustomed to this bizarre behavior and do not say what they are thinking. There is silence for a moment. The father continues: "Knife and fork aren't, but spoon is." The whine of the wheels and some labored breathing are the only audible sounds. "Autumn, winter, and spring aren't, but summer is. Breakfast isn't, and neither is lunch, but dinner is. Hammer is, but wrench, pliers, and saws are not. In fact, as you suspected, dear children, brothers and sisters aren't but mommies and daddies are. What do you think?"

Again, a relative silence reigned. Three minds had a thought that they did not express. The older child was a delicious little girl and knew better than to say anything about her father's sanity. "Are hot dogs?," she asked. "Oh, no," replied the father, with a twinkle in his voice. The younger child, a boy often dazzled by the three older members of his immediate family, asked, "What about pizza?" "Oh, pizza is," said the father.

Many question and answers followed. The wife fell asleep, but the children remained alert. Shortly before they pulled into a roadside restaurant for lunch, the little girl said, "Home isn't, the library isn't, the supermarket isn't, and the flower stand isn't either, but school is." It is sad to report that the father nearly lost control of the car, so overpowering was his glee, which of course 'is.'"

Children can find patterns earlier in their lives than many people think. Practice in finding those patterns helps to prepare them for more difficult explorations later in their lives. Read on.

GAME

2.2. *An Evil Puzzle: "Jeanette and Giuseppe"*

It was late on a stormy Sunday night, and the five young men at a nameless prestigious college were entertaining themselves as best they could in each other's company. Puzzles and problems flew and became increasingly bizarre. Suddenly, Igor's eyes flashed. "Have any of you heard about Jeanette and Giuseppe?"

Neither Bob, Chang, Mohammed, nor Ziggy had. Igor continued, "I am going to utter a phrase, after which you may ask as many yes-or-no questions as you wish, each of which I will answer. When you think you have solved the puzzle, touch your right middle finger to your nose. I will then ask you five yes-or-no questions, by which I will determine whether you have in verity solved the puzzle. The phrase is: 'Jeanette and Giuseppe.' "

"Is that a phrase?" asked Mohammed.

"Yes," intoned Igor, in his best pedantic manner.

"He's playing the game already," said Bob.

"That is not a yes-or-no question," continued Igor.

The questions and answers followed quickly.

Question	*Answer*
Are Jeanette and Giuseppe lovers?	No
Is Jeanette French?	No
Is Giuseppe Italian?	No
Are Jeanette and Giuseppe people?	Yes
Is Jeanette profligate?	Yes
Is Giuseppe a monk?	No
Are Jeanette and Giuseppe minions of Dracula?	Yes
Is Jeanette's number 666?	No
Have Jeanette and Giuseppe ever appeared in a Hollywood movie?	Yes
Am I going to pass Mathematics 1A?	Yes

Bob touched the middle finger of his right hand to his nose. Igor began the fatal questioning.

Igor's Questions	*Bob's Answers*
Are you contemplating dropping out of college?	No
Is your sister contemplating dropping out of college?	No
Is your mother contemplating going for her Ph.D.?	No
Is your Uncle Alexander going to get out of prison before the turn of the century?	No
Are you a roaring jackass?	Yes

Igor put the little finger of his left hand in his left ear. "This means that this is a comment outside of the game. Your last answer transcended the game and was right in a metaphysical sense." Igor removed his pinky finger from his ear. "You have not solved the problem," he intoned. "I will repeat the problem."

Igor repeated the entire spiel quoted before. The game proceeded, not for hours but for days, interrupted only by the necessities of college life.

Several weeks later, Ziggy and Chang raced across the yard outside their dorm and tackled Igor, throwing him on his back on the moist grass. He was not at all surprised to see each of them touching the middle fingers of their respective right hands to their respective noses.

Igor's Questions	*Unison Responses*
Does 2 times 3 equal 6?	No
Does 1 plus 2 equal 9?	Yes
Are the uses of adversity sweet?	No
Does music have charms to soothe the savage breast?	No
Does the King James edition of the Bible appear in the Congressional library card kat'l-og'?	It's not at all clear

Igor grinned. "You guys got it. I'm really impressed."

Are you, dear reader? . . . Read on.

⑥ 2.3. *The Josephus Problem*

⑫ NUMB One thousand persons stood in a (large) circle. Each, starting with the bald-headed woman, wore a sign on his or her back with a numeral from 1 to 1,000 in a clockwise sequence. They began counting off. The bald-headed woman said, "One, in," and remained in the circle. The 7′6″ basketball player to her right (they all faced out) said, "Two, out," and left the circle. The aesthetically repulsive Indic Philology major (of undetermined gender) next in sequence said, "Three, in," and remained in the circle. The cheerleader next in sequence (a lithe, attractive blond male) said, "Four, out," and left the circle.

So it continued, with each person sporting an odd numeral stating the numeral, saying "in," and remaining in the circle and with each person wearing an even numeral stating the numeral, saying "out" and leaving the circle.

It is easy to visualize who was left when the count-off once again reached the bald-headed woman. Since the person before her second turn had just said, "One thousand, out," and left the circle, she now said, "One, in," and remained in the circle. Continuing the alternating sequence, the Indic Philology major now said, "Three, out," and left the circle. This time, only those persons whose number was of the form $4n + 1$, n a positive integer, remained in the circle, while those whose number was of the form $4n + 3$, n a positive integer, left the circle. As the count approaches the bald-headed woman, we heard, "997, in," "999, out," followed by her intonation, "One, in." The next statement, by our previously encountered friend Igor (see section 2.2) was, "Five, out."

If this pattern of counting continues until exactly one person remains, you are asked to predict the numeral on that person's back.

It is certainly possible to program a computer to give us the answer to this problem, but we wouldn't gain any appreciation of the mathematics involved. You are therefore forbidden to use electronic devices or, even worse, trial and error. Any reader who writes the numerals from 1 to 1,000 in a large circle on a piece of oaktag and attempts, perhaps with many colored pens, to simulate the count-off process, is advised to close this book now and never reopen it.

How might we uncover the heart of this artichoke? In section four of this chapter, the Josephus problem will be revisited. Let it remain temporarily a morsel awaiting the lemon butter sauce of an ingenious solution. Read on.

NUMB 2.4. *The Long Cell Block*

An ancient prison was built in a long, long line, with numbered cells on one side of a stone corridor. One night, the jailer imbibed too freely, and as each prisoner slept, he unlocked every cell, walked back to cell 1, left it unlocked, but locked cell 2 and each subsequent even-numbered cell. Back he walked to the beginning, and proceeded as follows: He approached each third cell beginning with cell 3, and changed its state, locking it if he found it unlocked (as he did with cell 3) and unlocking it if he found it locked (as he did with cell 6). When he finished, he staggered back to the beginning and proceeded to change the state of every fourth cell, beginning with cell 4. When he had completed that tour, he went back to the beginning and repeated the process of changing the state of each fifth cell, beginning with cell 5. Each time he reached the end of the corridor, whether or not he was able to continue the pattern of that tour in which he was changing the state of each n^{th} cell, beginning with cell n, he returned to the beginning and changed the state of each $(n + 1)^{st}$ cell, beginning with cell $(n + 1)$. He concluded this intoxicated operation by returning to cell 1, walking the entire corridor, changing the state of the last cell, and collapsing.

Which prisoners were able to escape in the morning?

2.5. *Definitions and Discussion*

⑥
⑬
GEOM
NUMB

A recurring need—not only in mathematical research but in investigations in almost every academic and real-life area—is the ability to recognize, identify, and describe patterns. Patterns may be simple enough to become the crux of a childhood game such as "Moon Is, Sun Isn't," or may be subtle enough to elude the most brilliant minds of each generation, as was the proof of Fermat's Last Theorem. Developing skill at searching out patterns should be a major thrust of education at home, in school, and on the job. The beauty of a game such as "Moon Is, Sun Isn't" is that it models a form of pattern seeking that translates well into classroom instruction.

Suppose, for example, a teacher wished to introduce the concept of a parallelogram. The teacher might well begin by providing students with a picture sheet of perhaps 20 plane figures, some identified with the label *is* and the others identified with the words *is not*. At the first level, students might be asked to examine that page and create a page with 10 figures that they would label with the word *is* and another page with 10 figures which they would label with the words *is not*. On a third page, students would be asked to explain in their own words the criteria they used to decide which of their figures "is" and which of their figures "is not," and would then be asked to list, to the best of their ability, the clues from the teacher's original page that led to their stated decision criteria. Students in small groups would discuss the three pages and attempt to arrive at a simple and clear statement of the decision criteria upon which they can agree. They might then be asked to create an appropriate word for the figures listed under "is." Many students will come to a good definition for what we call a parallelogram and will observe the need to prove or disprove that criteria established by different groups are equivalent.

Thus, for example, Group A might say that the figures listed as "is" had opposite sides equal in length. Group B might say that the figures were those that had opposite angles equal in measure. Improbably, but possibly, Group C might say the "is" figures were those whose diagonals bisected each other. Probably one or more groups would identify the "is" figures as those that had opposite sides parallel. The teacher would facilitate the development of a sound definition by introducing the need to replace the word *figures* by *quadrilaterals* and by suggesting that a powerful line of investigation would be to see whether the conditions that each group found could each be proved from others of the listed conditions. A final outcome might be a powerful definition of parallelogram as a quadrilateral that has any one of, and hence (as proved before the conclusion of the unit) all of, the properties listed below:

(The list is left to the reader.)

In the game "Moon Is, Sun Isn't," a newly literate child might not write a long composition but would do well to note, even if full spelling were not yet within the child's capability, that each *is* word had a double letter (consecutive), whereas, the *isn't* words did not.

Jeanette and Giuseppe is a hidden-pattern game suitable for older children and adults. It requires considerable sophistication to realize that the answers to the questions have nothing to do with the content of these questions. The pattern illustrated in Section B was simply that each question whose final word ended in vowel was answered "yes"; each question whose final word ended in a consonant was answered "no." The uncertainty related to the question in the section in which the final word was given phonetically as "kat'l-og'" is that both *catalog* and *catalogue* are acceptable spellings in English. This game develops flexibility in pattern-seeking mindsets by giving students the confidence to act on seeming contradictions. When the answers to their questions seem to be inconsistent, they must move beyond the urge to give up to the urge to look at the problem from an alternative point of view (a strategy we explore more fully in the next section).

The Josephus problem will not be discussed in this section, because it yields to a combination of problem-exploration strategies. It will be analyzed in section 4.2.

The Long Cell Block problem is suitable for students cognizant of the possibility of factoring integers uniquely (except for order) into primes. By making factorization tables after coming to the realization that the state of a cell is changed once for every factor including one, they will see that for most positive integers, factors come in distinct pairs. Thus, for 6, the pairs are respectively (1, 6) and (2, 3). Cell 6 will be opened for 1, closed for 2, opened for 3, and closed for 6; in the morning, it will be locked. Some students will come to the exciting discovery that the only positive integers with an odd number of factors are those in which one of the factors is paired with itself in the factorization, so that it leads to a *single* change of state. Clearly, these special positive integers are the perfect squares. Thus, for example, the pairings for 9 are (1, 9) and (3, 3). The pair (1, 9) leads to the expected two changes, but the pair (3, 3) leads to only one change. Thus, cells 1, 4, 9, 16, 25, and indeed all cells in the block of the form n^2, n a positive integer, will be open in the morning. Most people learn best by constructing their own knowledge. The samples given here model that constructivist philosophy using words, number patterns, and, in the case of the described activity with parallelograms, pictures. There is no need to give students a formal definition of *pattern* or *finding a pattern*.

With enough suitable practice, almost every student will come to preverbal understanding, and once that preverbal understanding has been reached, the student will grow in the ability to apply the problem-exploration strategy of finding patterns.

3. Adopting a Different Point of View

3.1. An Ancient Chinese Fable: The Men and the Wall

GEN

A good illustration of differences between problem solving and problem exploration comes from an ancient Chinese folk tale.

Four men are walking across a field when they come to a great wall with a massive gate locked by a thick chain and padlock. Three of the men attempt to force or break the padlock, using rocks, logs, and fire, to no avail. The fourth man walked into the nearby woods. Twenty minutes later, the three problem solvers were tired and frustrated. Suddenly, they saw the fourth man running out of the woods toward them holding a long shoot of bamboo. He planted the pole in the earth as he approached the wall at high speed and lightly vaulted over it. He had not solved the problem of opening the lock; he had redefined the problem. Having seen no solution to the unstated problem of opening the gate, he created the new problem of getting to the other side of the wall.

His solution was elegant and simple; it was not final and conclusive. Two of the remaining walkers were too old and too feeble to vault the wall successfully. Was the new problem to find a way to get *them* to the far side of the gate?

The problem explorer asked his companions to pass the pole through the gate to him. He retreated, ran toward them, and vaulted back. "Where should we walk instead?" he asked. The fourth man said, "I would love to know where this wall ends and see if we can find out why it was built. Let us continue our walk in its shadow. Perhaps we will meet someone who can tell us the reason for the wall."

Problem explorers are not limited by stated problems or by common expectations. They see ill-structured problem situations as a gestalt. They ask many questions and pursue those for which attacks can be devised. Often, they redefine a stated problem into a more productive and accessible form.

3.2. Sharing Pudding Fairly

ARITH

GEN

Two gluttonous children wish to share a bowl of pudding. They want as much as they can get but eschew physical combat. Having learned hypocrisy from their elders, they purport to seek fairness in the division of the pudding.

Having heard the word *fair*, two nearby adult butters-in approach the children with the usual adult closed solution. The first adult, Ann, suggests, "Let us get a balance scale and two plates. We will divide the pudding so that the two pans are exactly even, and thus ensure that each of you gets a fair share."

Ann did what she had promised. To her practiced eyed, the two pans were perfectly even and justice was guaranteed. Gluttonous Suzie began to whine in her accustomed way, "Johnny has more than I do."

"We haven't said which dish is Johnny's and which is yours, dear," Ann offered.

"It doesn't matter," screamed Suzie, "whichever one you give him, he'll have more than I have."

Counselor Bob had been standing by silently. "Fairness," said Bob pedantically, "is psychological, not scientific. You do everything too scientifically, Ann, and I should know. Dear children, let me tell you what to do: Let's dump both plates back in the bowl. Suzie, you divide the pudding in a way you think is fair. Johnny, you select which plate you want after Suzie finishes. You see, Suzie, if you divide the pudding fairly, whichever one Johnny picks will be OK. And Johnny, you pick whichever you think is more."

"That's no good," Johnny shouted, "I want to divide and let Suzie pick."

The problem explorer, contemplating the ill-structured situation just described, recognizes a multitude of implicit problems. Perhaps the explorer decides that the real problem is to divert the children from the impending war. King Solomon might say, "Divvy it up any way you want, and when you're finished, if I don't hear any screaming or fighting, you can each have a banana split with three scoops of ice cream and ample hot fudge sauce."

A nearby mathematics professor, Carl Friedrich, might abstain from any contact with the people involved and pursue the following thinking: "I should be able to generalize this problem to *n* persons, using the scientific approach with no difficulty. How can I generalize Counselor Bob's psychological approach to three people, let alone to *n*?"

Fairness is a deep concept, whether it is viewed mathematically, psychologically, socially, politically, or humorously. Wonderful and rich interdisciplinary instruction in schools, from kindergarten through graduate school, could center on fairness. Redefining *fairness* from a scientific to a psychological context is an insightful shift that might elude many closed-minded, closed-ended problem solvers.

ARITH

GEN

3.3. *A Talmudic Tale: The Men, the Loaves, and the Gold*

On a *stetl* farm in Poland many years ago, three men were working in the field as the time for the midday meal drew near. One of the men, Ezekiel, was the wealthy owner of the field who believed that by working one day each year with the men who lived in poverty on his land, he would achieve the Lord's blessing and the men's loyalty. He noted that he had 10 gold coins in his pocket, but that the workers had brought loaves of bread. Abraham had three loaves and Isaac had two.

"I have a splendid idea," Ezekiel proudly declared. "Let us share the five loaves you have between you equally among us, and in return I will give the 10 golden coins in my pocket to share between the two of you." Abraham and Isaac were overjoyed, recognizing that a single gold coin was nearly a year's income. The three men shared the loaves. After eating, Ezekiel announced that he had grown fatigued and would cease work for the day, and that his men could also quit work early at sundown.

As soon as Ezekiel left, Abraham and Isaac began to argue. Abraham said, "I had three loaves and you had two. I should get six gold coins and you should get four, right?"

"No, no," said Isaac. "We shared the bread equally, we should share the gold equally, five coins apiece."

Long before sundown, the two men decided to refer their dispute to their rabbi, who was noted not only for his talmudic wisdom but for his mathematical acumen.

"So, Rabbi Jacob," Abraham concluded after recounting the events of the late morning, "what should we do?"

A furtive tear appeared in the corner of Rabbi Jacob's eye. "Dear friends," he said, "you present me with a terrible problem. If on the one hand I advise you to share the coins as Abraham wants, you, Isaac, will be a trifle aggravated, but you will soon concur and you both will leave me to my studies. If I counsel you to divide the coins five and five, you, Abraham, will be annoyed, but soon you will concur, and, again, both of you will leave me to my studies. On the third hand, unfortunately, if I tell you the correct solution, both of you will argue with me until dawn, I will not get back to my studies for a day, and neither of you will be fit for work tomorrow."

In chorus, Abraham and Isaac shouted, "What do you mean, Rabbi Jacob? Surely, there are no other possibilities."

Rabbi Jacob, who had foreseen section 3.2 about pudding apportionment those many centuries before it was written, was faced with a quandary. He was obligated to tell the truth by his own vision of rightness and rabbinical responsibility. He knew that this would not go well.

"My friends, I ask you to consider the reason that money is paid for goods. The gold is intended to compensate the seller for what he has given to the buyer in goods and in service. In this case, we need only consider goods. Abraham, you and Isaac each ate one and two thirds loaves of bread, as did Reb Ezekiel. Therefore, you, Abraham, gave away one and one third loaves of bread to your honored landlord; whereas, you, Isaac, gave away only one third of a loaf of bread. Since you, Abraham, gave away four times as much bread as Isaac, you should get four times the compensation. The fair distribution of the 10 gold coins is 8 coins to Abraham and 2 coins to Isaac."

The screaming continued all through the night, and Rabbi Jacob did not get back to his studies for several days.

② ⑥ ARITH

3.4. Adding and Subtracting Signed Numbers

"I get most of it, Mrs. Thorn," Jacqueline said. "When you add two positive numbers, I understand why you add the values and affix a positive sign. When you add two negative numbers, I understand why you add their values and affix a negative sign. I begin to get confused when you talk about problems like 'positive seven plus negative three' as adding because, to me, it's subtracting. In fact, 'positive three plus negative seven' is a little worse. But I can get all the answers you want to all of those problems, even though I'm not sure what those answers mean. I really get into a lot of trouble with subtraction.

"When you give us 'positive seven minus positive three,' I get positive four, just like you want, but I have a lot of trouble when the second number is negative. For example, 'positive three minus negative four' makes no sense, no matter whether I give you your answer of positive seven or my answer of negative one. Y'know, Mrs. Thorn, I never even understood 'one minus three' back in grade school. If you put a penny on the table and tell me to take three pennies away, I can't owe the table two pennies. I really think that 'one minus three' has no answer."

Mrs. Thorn considered herself to be an exemplary teacher, cognizant of all the most modern methodologies, doctrines, and insights related to her noble calling. She took out a box of manipulatives and told Jacqueline to work with them until she understood perfectly all the concepts involved. Jacqueline knew exactly what to do. She puttered with the junk for 11 minutes and then announced joyously, "I get it, Mrs. Thorn." Since Jacqueline was quite proficient at memorizing rules, she knew that she could get high grades on all of Mrs. Thorn's tests and that her wonderful work with the manipulatives would earn her extra points on the report card.

"Nobody really understands this stuff," Jacqueline thought, "but then, nobody cares about mathematics anyway after you get out of school."

It is interesting to speculate how most teachers would respond to the question, "Does positive one equal one?" My observation has been that the vast majority of teachers I have met would say something meaning, "What a stupid question! Of course they're equal."

Are they really?

If I were to hold up two fingers and ask a student to state how many fingers were extended, the correct answer, in the deepest sense, is not positive two but two. Positive and negative numbers are measures of change, not of status. If I were to hold up two fingers, put down my hand, and then hold up four fingers, and then I were to ask a student for the number that measures the change from two to four, the correct answer is positive two and not two. Similarly, if I were to hold up four fingers, put down my hand, and then hold up two fingers, asking what number measures the change from four to two, the correct answer would be negative two, not two. The number two is not equal either to negative two or to positive two.

We are dealing with something more basic than semantics. Fundamental understanding of signed numbers requires far more than knowing how to get the answer that the teacher wants. The meaning of "positive two plus positive three" is very different from the meaning of "two plus three." The elementary arithmetic problem of computing two plus three is properly modeled by pouring three marbles from a glass into a glass containing two marbles and counting how many marbles are in the glass at the end. An appropriate model for positive two plus positive three might be: "Little Mary gets a two-cents-a-week increase in her allowance in March. In May, she gets a three-cents-a-week increase. What is the change in her allowance from the beginning of March to the end of May?" Clearly, the correct answer is that she had a five-cents-a-week increase and not simply five cents.

In fact, if a child were asked to compute one minus three, the very best answer the child could give would be as follows: "Technically, the correct answer to your problem is that it cannot be done. Perhaps you want me to embed the natural numbers isomorphically into the reals and give you the answer negative two; I think that's a bit advanced for fourth grade."

Perhaps after all I will never hear that answer. I would be happier hearing "It can't be done" than hearing "negative two."

Jacqueline would be much closer to nirvana if she had been taught to think about the *meaning* of "positive three minus negative four" rather than the answer. The conceptualization of subtraction of signed numbers should not be dependent on an abstract mathematical concept such as adding an additive inverse.

Change the point of view. There are several models for "seven minus three" in arithmetic. There is the "take away" model: "We have seven marbles, we take three away, how many are left?" There is the "buildup" model: "What do we have to add to three to get seven?" There is the "number ray" model: "How far is it between the point marked 'three' on our number ray and the point marked 'seven'?"

When expanding student conceptualization of number systems from whole numbers to integers, it would be most appropriate for the student to examine each of the three arithmetic models and to decide which of the three most readily yields insight into subtraction of signed numbers.

The student might think as follows about positive three minus negative four:

1. I have an increase of three and I take away a decrease of four. Hmmmm. I might be able to make sense of this in a few months, but I'd rather not try now.
2. What must I add to negative four to get positive three? Hmmmm. If I have a decrease of four, I need an increase of seven to achieve a total increase of three. I guess positive three minus negative four should equal positive seven.
3. When we work with integers instead of whole number, I guess we replace the number ray with a number line. Positive three minus negative four

> would now mean, "What number measures the change needed to get from the point marked negative four to the point marked positive three on the number line?" I see you'd have to go seven units *to the right*, so the answer must be positive seven.

The point is that looking at subtraction of signed numbers from several perspectives arising from appropriate models for arithmetic subtraction, the student can obtain direct and lasting insight into fundamental meanings without memorizing partially understood rules. As a bonus, the student begins to think along lines that might one day help him to construct meaning for other generalizations such as negative or fractional exponents, trigonometric functions of rotations greater than 90 degrees or less than zero degrees, or the extension of the factorial function to the gamma function.

②

NUMB

PROB

3.5. *Discussion Without Definition*

Patterns are omnipresent but not always easily discernible. The investigations discussed in this section have explored patterns in contexts that are too infrequently brought into the classroom. The most familiar context for discussing patterns in mathematics has been through the use of sequences; this creates a very limited perception as to what shapes and forms patterns can assume.

Even when sequences are used to elucidate patterns, there is too much of a tendency to use recursive arithmetic operations to define a sequence. Only rarely do subtle and provocative samples of sequence patterns reach the classroom.

Here is a wonderful sequence pattern with a missing term to be determined and explained.

10, 11, 12, 13, 14, 15, 16, 17, 20, 22, 24, ___, 100, 121, 1,000.

The above sequence is complete and hence finite. The identity of the missing term and the explanation for the sequence are wonderful to contemplate. Were I to give an explanatory title to the puzzle, that title would be "Oh, So Sweet."

Two of the problems in this section dealt with fairness. The extension of the pudding sharing in a psychologically fair way to three gluttonous children can be done as follows. Let us call the children Dick, Jane, and Spot. The order in which the action to be described is taken is determined by flipping a fair three-sided coin or, if none is easily available, tossing a die, with Dick going first if ace or deuce comes up, Jane going first if trey or four is on top, and Spot getting first lick if five or six is uppermost.

We set a context in which Dick goes first. Dick creates a portion of pudding that he thinks is fair—that is, a portion he would gladly accept.

Jane and Spot proceed sequentially. If Jane believes that Dick's portion is a fair share or less, she passes, and it is Spot's turn to decide. If Jane believes that the portion Dick created was too large, she stakes a claim by returning some pudding to the big bowl and saying that she would take the reduced portion as her own.

Ostensibly, this would be acceptable to Dick, since Jane is taking a portion smaller than one which he considered fair. It is now Spot's turn. Spot can either agree that the portion Jane is holding is to be hers or may stake a final claim to it by removing some of the pudding and returning the removed pudding to the big bowl, making the portion Spot's and Spot's alone. Once any one of the three gluttons has been awarded a portion by this procedure, the remaining two partitioned the pudding in the big bowl by the original two-person technique: "one divides, the other chooses." (Preferable designation would be "one apportions, the other selects," because *divide* and *apportion* should not be used synonymously in mathematics class. To see why, think about dividing six into two parts as contrasted to apportioning six marbles among two students.) The generalization of the psychologically fair apportionment technique to four or more gluttons is left to the reader. I hope the reader thinks this is fair.

The advantage of finding a pattern while exploring a problem situation is that much broader predictions can be made about the structure in which a pattern exits. In extreme conditions, the exact nature of an unfamiliar structure is revealed. Consider the finite sequence mentioned above. If the listed numerals are viewed conventionally, little is discernible. Suppose, on the other hand, that a practiced problem explorer viewed the 15 numerals (actually, 14 numerals and a blank) as each representing the same number. Perhaps it would be a deft leap to the vista from which can be seen that the number being named is 16 (oh, how sweet) and that the bases of numeration are, in order, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2. It is undoubtedly unnecessary to inform the reader that the numeral to be inserted in the blank is 31.

3.6. *Changing Points of View About Patterns*

② ⑨ ⑬ ALG NUMB GEN

The description of a discovered pattern may take several forms. If the pattern is a sequence or a linear geometric pattern, there are two simple ways to describe the pattern: explicitly or recursively. An *explicit* description of a sequential pattern is a formula or complete description that gives the term or figure in n^{th} position as a function of n. Thus, for example, an explicit description of the arithmetic sequence $\{ a_n \}$ with initial term $a_1 = 3$ and with common difference $d = 4$ is $a_n = 3 + (n - 1) \times 4$. This means that the seventh term in the sequence is $3 + (7 - 1) \times 4 = 3 + 6 \times 4 = 27$, and that any term can be immediately calculated.

A *recursive* description of the pattern of the sequence is one that gives the n^{th} term in terms of a finite number of previous terms. Thus, a recursive description of the arithmetic sequence just discussed would be given by $a_n = a_{n-1} + 4$. This description tells us that, no matter where we are in the sequence, we can get the next term by adding 4 to the term we have.

Explicit and recursive descriptions of linear geometric patterns are also possible. For example, a long sequence of squares and circles that alternate, beginning with a square, could be described explicitly for n a positive integer by a_{2n-1} is a one-by-one square with sides parallel and perpendicular to the linear sequence of shapes, a_{2n} is a circle of diameter one with top and bottom such that if the circle were moved to the left or right, the circle would be tangent to the upper and lower bases of the square. A recursive description of the same linear geometric pattern would be that the pattern consists solely of one-by-one squares and diameter-one circles oriented so that the sides of the square are parallel or perpendicular to the pattern and that the circle, if moved into an adjacent square, would be tangent to the upper and lower bases and such that a_n is a circle if and only if a_{n-1} is a square. In both cases, the sequence and the linear pattern, the recursive definitions just given are incomplete, because we were not told how to begin. We must append to the recursive definition of the numerical sequence that $a_1 = 3$ and to the recursive definition of the geometric sequence that a_1 is a square.

Computers can handle both recursive and explicit definitions of patterns effectively. Mathematically, we can pose fascinating problems by giving an explicit description of a sequence and asking for a recursive one, or giving a recursive description of a sequence and asking for an explicit one. Thus, for example, the classical Fibonacci sequence may be defined recursively by $a_1 = a_2 = 1$ and, for $n > 3, a_n = a_{n-1} + a_{n-2}$. Given that definition, it is an interesting challenge to deduce an explicit description. It is interesting to speculate whether the recursive or the explicit description of the Fibonacci sequence gives greater insight into its nature and when each is preferable for further investigation of the sequence.

Both recursive and explicit descriptions are possible for higher-dimensional numerical and geometric patterns. Consider whether one would be more likely to obtain the designer's view of a large tesselated piazza if one were to present the artisan with explicit instructions for which tile to place on each coordinate of a design grid or by giving the artisan a rectangular block containing the complete design with the instruction to place that block in the upper left-hand corner and to duplicate exactly that design horizontally and vertically throughout the (presumably) rectangular piazza, with no interstices and with parallel orientation of all edges.

4. Solving a Simpler Analogous Problem

GEN

4.1. *The Budget*

Professor Waarwoolf of Transylvania Tech was attempting to teach his graduate business seminar the intricate details of creating an annual budget. His model problem dealt with American Megatech, a highly diversified corporation with tentacles in every cranny of the civilized world. On the third day of his intensive lectures, Professor Waarwoolf was dismayed to note that a hand was raised toward the middle of the room.

"I'm not sure I'm getting this," said the soon-to-be-former graduate student. "Might we not address the same issues just by examining proper allocation of weekly allowance funds made by my little brother Vlad?"

A chorus of odd affirmative snorts was heard throughout the lecture room.

Professor Waarwoolf had fallen into an honored and ancient pedagogical trap. By seeking generality, he had achieved obfuscation. A student who thoroughly understood every ramification of Vlad's weekly finances would be better able to ask appropriate questions and organize the insights necessary to expand from Vlad to the agrarian collective to the state budget and, ultimately, to American Megatech.

4.2. *Reexamining Josephus*

② ⑥ ⑪ NUMB

Let us return to the thrilling lines of yestersection 2.3. Polly Sue Seidlitz had encountered the Josephus problem in her sixth-grade mathematics problem-exploration club. She had been thinking about it for several days but could not bring herself to write the numerals from 1 to 1,000 on a piece of paper.

"Teachers are so inconsiderate," she thought. "If Mr. Sperdel had made it 10 people instead of 1,000, I could have done it."

Polly Sue hoped she was not deceiving herself. She wrote the numerals from 1 to 10 in an appropriate circle on the inviting blank sheet before her. She picked up her red pencil and crossed out 2, 4, 6, 8, and 10. She put down the red pencil and picked up the blue one, thinking, "1 in, 3 out, 5 in, 7 out, 9 in." She crossed out 3 and 7 in blue. It was time for the green pencil. "One out, 5 in, 9 out" resulted in appropriate crossouts. One numeral remained. "I could have solved it if he had said 10! The answer is 5." She stared at her work for a long time.

A spark within her spirit burst into flame. "I'll bet that if I solve the problem for enough small numbers, I'll see a pattern."

Polly Sue took another blank sheet on which she began the following chart:

Number in the Circle (N.C.)	*Number on Back of Last Remaining Person (N.B.L.R.P.)*
1	1
2	1
3	3
4	1
5	3
6	5
7	7

Sally stared at the chart. "I'm tempted to write '8, 9' on the next line, but there is no number 9 when there are only 8 people, so what sense would that make? I'll bet the answer turns out to be 1 again." A little colored-pencil play verified her conjecture.

Polly Sue contemplated the chart. "I'll bet the sequence of odd numbers starts anew. After all, it would figure, knowing that the answer for 10 is 5, that the answer for 9 would be 3." More colored-pencil play resulted in the following chart:

N.C.	*N.B.L.R.P.*	*N.C.*	*N.B.L.R.P.*
1	1	11	7
2	1	12	9
3	3	13	11
4	1	14	13
5	3	15	15
6	5	16	1
7	7	17	3
8	1	18	5
9	3	19	7
10	5	20	9

Polly Sue went to the bookcase and took down her father's copy of the book you now hold in your hand. She opened to section 3.6 and read it with astounding comprehension for one so chronologically young. "I can continue this chart recursively with no trouble. I wonder if I can get an explicit formulation that will let me express the right-hand term on any row, knowing the left-hand term.

Polly Sue re-examined the rows in which the right hand term was 1. "What would Daddy say . . . I mean, what could I say is common to 1, 2, 4, 8, and 16?" It was crystal clear. "Powers of two!" Polly yelled with the enthusiasm with which many of the boys in her class yelled "Boom" when they threw water bags at her. She filled in her chart up to 32, following the indicated recursive pattern and ending with:

N.C.	*N.B.L.R.P.*
30	29
31	31
32	1

"Let's see if I can remember . . . uh, reconstruct what Daddy made me do when he locked me in the room last August until I came up with general algebraic forms for odd and even numbers. Oh, yes, the even numbers were of the form $2n$, n a positive integer (in this case), and the odd numbers were of the form $2n + 1$ or $2n - 1$."

Much writing and some scribbling resulted in the following conclusion:

For n a whole number, if the number of people in the original circle is of the form $2^k + n$, then the last person in the circle will have the numeral for $2n + 1$ on his back.

"I can solve this problem. Let me write the consecutive powers of 2: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 . . . mustn't pass 1,000. Since $1{,}000 = 512 + 488$, the numeral on the last guy's back will be $2(488) + 1$ or 977. Boy, I never would have guessed that."

4.3. *Adding Fractions*

ARITH

Mr. Sperling had had a very difficult morning. He had been issued a speeding ticket on the way to school and was brooding about it, and his fifth-grade class was treating him gingerly; they recognized storm signals.

"Today we are going to learn how to add fractions," barked Mr. Sperling. "Try to add 3/17 and 6/13." There was dead silence. Even Jerry, one of the very best math students Mr. Sperling had ever taught, stared at the blackboard with puzzlement. Mr. Sperling already regretted having taken out his traffic ticket on the class, but he knew no good way to get out of the bind.

Jerry wrote something on his paper, studied it for a few minutes, and then raised his hand.

"Mr. Sperling, those numbers are too big for me. I always try to learn something with smaller numbers because if get it with the small numbers, I usually can get it with the big ones. I tried 'a half plus a half' and thought like this: One of anything added to one of the same thing gives two of that thing. So a half plus a half equals two halves and two halves is one whole. I'm not sure why, but I don't feel good about saying it that way. You [Mr. Sperling] never say 'One half plus one half equals one unit,' but somehow I see 'unit' as an appropriate abstraction and I don't see 'whole' that way." Jerry fidgeted a little more. "I really don't have it yet because thinking through a half plus a half just doesn't do it. It's a bit scary, but I've got to think about a half plus a third. Can I go over into the math corner for a while?"

A glimmer of a glimmer of a glimmer of a perception came to life in Mr. Sperling's gray cell morass. There was a spiritual bond between Mr. Sperling and a German teacher from a couple of centuries ago who respected the thinking of an odd child in his class who found an unusual way of adding the numbers from 1 to 100 mentally. Unarticulated but present was the thought that by nurturing this Jerry-thing he might one day be celebrated.

"Do you see what wonderful thing Jerry is doing?" said Mr. Sperling to a class of students who were too stunned by what was happening to do anything except stare with glazed eyes at a fixed point at infinity that they had never previously contemplated. "I purposely gave you large numbers to add, hoping that you would learn the wonderful strategy of simplifying. Many times you can gain deep understanding of a problem situation you encounter by addressing a simpler, analogous problem situation."

"Ann *who?*" asked Betty, Mr. Sperling's front row orchid.

Ignoring her honest inquiry in pursuit of knowledge, Mr. Sperling continued, "Jerry saw that the tops of fraction are called *numerators* because they tell us how much or how many (the *number*) we have of the unit named in the bottom, which we call the *denominator* (the *name*). Jerry saw that thinking about one half plus one half was an analogy (remember our lessons on analogy, Betty?) to a problem such as adding one dog to one dog and getting two dogs. Jerry is off in the math corner trying to figure out what one dog plus one cat equals."

Jerry started to laugh. He spoke without raising his hand or gaining the traditional license to speak. "I think I understand what adding fractions means, Mr. Sperling. You know, you have been wrong every time you told us that you can't add apples and oranges. Of course we can add apples and oranges. One apple plus one orange equals two pieces of fruit. One dog plus one cat equals two animals or two mammals or two living things. All we need do is find a unit which takes in both cats and dogs or, in the other case, both apples and oranges. So I said to myself, 'Hey Jerry, what takes in both halves and thirds?' So I said, 'Certainly not fourths or fifths, but sixths, twelfths, or twenty-fourths would be just fine.' Well, Mr. Sperling, it's really very easy. One half is three sixths and one third is two sixths, and I have reduced the problem of adding halves and thirds to the simpler and previously solved problem of adding two numerical expressions with identical units. Two of anything plus three of the same thing is five of that thing."

There was a gentle knock on the door. Polly Sue Seidlitz entered the room with a note from Mr. Grimmas, the principal, to Mr. Sperling about the traffic officer waiting in the general office. As Mr. Sperling read the note and became appropriately ashen, Polly Sue glanced at the blackboard and saw the original problem of adding 3/17 and 6/13. "How can fifth grade be doing that when our sixth-grade Bluebird class never does things this hard?"

"Take over the class, Polly Sue, I've got to go to the principal's office," whined Mr. Sperling as he scooted out the door.

Polly Sue had waited all her life for the star to break his leg. "What's your thinking about this, class?" she asked, without knowing what "this" was. "Are you Polly Sue Seidlitz?" asked Jerry, with far more respect than he had ever accorded Mr. Sperling. "I've heard about you. I can explain the problem on the board."

She glanced at Mr. Sperling's oaktag with the kids' names on it. No one sat in the corner of the room with the polygons hanging on the wall. "What's your name, little boy?" she asked in as kindly a manner as such a child could manage.

"Jerry McBourbak," he answered.

"Well, Jerry, tell us about your thinking."

"In order to analyze this unfamiliar problem with big numbers, I first looked at a simpler problem like it, namely a half plus a third. It was a question of getting a common unit for halves and thirds, which could be sixths, twelfths, twenty-fourths, or so on."

"Or eighteenths," said Polly Sue.

"Hey, you're right!" cried Jerry, with genuine enthusiasm. "I missed that."

"Actually, so did I the first time," admitted Polly Sue. "Go on with your thinking."

"Well, with seventeenths and thirteenths the arithmetic was little harder, but the idea was exactly the same. I multiplied 17 by 13 in my head, using the 15 times 15 minus 4 model—I'm sure you know what I'm talking about, Polly Sue—and since I memorized all perfect squares from 1 × 1 to 31 × 31, I knew that 15 × 15 is 225, so 17 × 13 is 221. So the right unit for the answer could be 221sts or other stuff I'm not going to bother to figure out. So 3/17 is 39/221 and 6/13 is 102/221. Clearly, 3/17 plus 6/13 is 141/221. Apples and oranges," Jerry concluded.

"One man, two automobile tires, and the four laws of Kepler," said Polly Sue, smiling at a joke that was not a joke.

Many years later, as Jerry and Polly Sue lolled on the beach in Waikiki on their honeymoon, Polly Sue said, "We really should resolve emerging problems in our relationship before they fester. I never fully internalized the analogy between adding halves and thirds by combining the new unit sixths, changing the count of halves and thirds to conform to the new unit, to adding one apple to one orange and getting two pieces of fruit without having to change the count."

Jerry held her tenderly and explained with no hint of patronization what he had realized in a flashback in the math corner so long before. "Sometimes the count will change and sometimes it will not when you change units. To mail a letter requiring $1.20 postage, the current going rate for a one-ounce local letter" [after all, a couple of years have gone by] "you can paste on one $1.20 foil stamp, three 40-cent disks, six 20-cent rectangular traditionals, or 120 green antiques. Here, the count changes. On the other hand, one dog is one animal is one mammal is millions of chains of DNA bound in a complex pattern. Perhaps if we added one apple plus one orange plus 17 cells of a pear, the analogy would be clearer."

"Again and again you show me why I love you," said Polly Sue, author of the recent smash best-seller *Infinite Dimensional Fractal Manifolds and Their Measures.*

4.4. Finding the Centroid of a Polygon

GEOM

Polly Sue Seidlitz and Jerry had taken to doing their homework together. He often let her carry his books, because they both agreed that she was physically much stronger. They were going over to her house that afternoon because they had a particularly delicious set of common assignments for their problem-exploration club at Gallifrey High School. "I think the Doctor is one of the best teachers I've ever had. I got a kick out of his asking me to find the center of a 6-7-8 triangle." Polly Sue was exuberant.

"There are so many ways to think of center, even in a triangle," Jerry noted. "When I was five years old, Mom bought me a compass."

"You've got to learn to be more precise," Polly Sue exhorted. "I'm not even sure what you mean; something to find the magnetic north pole, or something to draw circles?"

"Sorry. She bought me a pair of compasses. I tried drawing circles around everything I could find and then I tried putting circles inside of them. I found it hard to get a single circle to intersect the vertices of a given triangle and even harder to get a single circle internally tangent to the sides or sides extended of a triangle."

"So which did you decide was the center of the triangle? The center of the inscribed circle, or the center of the circumscribed circle?"

"Neither, both, and more," Jerry responded. "After I found out how to circumscribe a circle around a triangle, I cut cardboard triangles out of my father's shirt cardboards and tried to balance them on a pin at the center of the circle. When that usually didn't work, I tried the same gig with the center of the inscribed circle. It took a while for me to find out about medians and centroids."

"Now that you know how to find the centroid of a triangle, do you think that's the true center?"

"Center is as center does," said Jerry.

"Well," Polly Sue said, "even as the Doctor was asking us to think about what we meant by the center of a triangle, I was way ahead of him for a change. I asked myself how one would find the centroid of an *n*-sided convex polygon. Have you ever thought about that?"

"Often, when I'm alone. I actually think that the Doctor gave the problem to you in a weaker pedagogical sequence than he might have. Suppose *he* had asked you how to find the centroid of an *n*-sided convex polygon: Consider what you would have had to do!"

"I get it! We *all* would have had to think about solving the simplest case of the problem, namely, finding the centroid of a triangle. We would all have learned something about problem exploration in addition to learning about the problem itself."

"Exactly. After you had simplified your general question to one about triangles and had learned all about the centroids of triangles, you would have proceeded naturally to convex quadrilaterals. When you draw a diagonal in one of those suckers, you get two triangles and can easily determine the centroid of each. It follows that the centroid of the quadrilateral is on the line segment joining them."

"That's obvious, but where?"

"Clearly, not the midpoint, at least not in general. The only time I'd expect it to be the midpoint would be when the two triangles were equal in area. That certainly doesn't happen most of the time."

"Sowhatchagonnado?"

"What we always do."

"What's that?"

"You know," Jerry said, with the innocence of a young mathematician. "Draw the other diagonal."

"Oh, where the analogous segment meets the one you drew," Polly Sue exclaimed. "That's great. Let's try to generalize that to convex pentagons. Looks easy: Draw a diagonal and you have a quadrilateral and a triangle. We can find the centroid of each and therefore have a line segment which must contain the centroid of the pentagon."

They had reached the steps leading to her front porch and were both quite excited. "Then use another diagonal and it works out just as with the quadrilateral. That generalizes to any convex polygon by induction."

"Naturally," said Polly Sue, inserting her key into the waiting keyhole and proceeding to unlock the door.

"Sometime let's look at other kinds of centers for polygons when they are defined."

4.5. Discussion

Oftentimes, the complexity of a setting in which a structured problem is given or in which many ill-structured problem situations are inherent impedes internalization of central concepts. For that reason, problems based in the real world are often destructive to the formation of sound mathematical concepts. It may be a serious error to embed important problem situations in artificial real-world settings, trading achievement of deep insight for an easy answer to the often unimportant questions, "How will we use this?" and "What good is this?"

Rather than seeking complex real situations, experienced problem explorers often prefer simplifying problem situations to the extreme. Thus, for example, when Polly Sue and Jerry were asked to make generalizations about polygons, they naturally first explored the question with triangles. When the class was asked by Mr. Sperling to add 3/17 and 6/13, Jerry brilliantly examined not the given problem but the addition of 1/2 and 1/3. He saw that thinking about 1/2 plus 1/2 was too much of a simplification, because identical denominators were involved. In the Josephus problem, Polly Sue charted solutions beginning with 1 person in the circle and continued increasing the number in the circle until a pattern emerged. She did not attempt the analysis of the problem with 1,000 persons in a circle. Finally, Professor Waarwoolf's student insightfully suggested studying the budgeting of a small boy's allowance so that principles and relevant ideas might be examined without the extraneous complexities of looking at a diversified corporation.

This is not to suggest that proper teaching proceed from the simple to the complex. Such teaching may result in temporarily faster integration of new ideas by students but may well not be the best way for them to learn techniques for profitable and imaginative exploration of ill-structured problem situations. Teachers should often pose structured problems in complex settings when they are teaching students how to find simpler analogous problems. This technique is not always apparent to students; it can and should be taught. An excellent paradigm for instructing in this area is this: Proceed from the complex to the simple to the complex. Ideally, students should suggest that this be done even when teachers do not structure instruction in this way.

Of course, if real-world settings of problems are given precisely for the reason of teaching students to proceed from the complex to the simple to the complex, and methodologies for doing this have been thought through, a major objection to using the real world to teach mathematics is removed, whereas the benefits remain.

5. Considering Extreme Cases

②

GEOM

5.1. *The Area Enclosed by a Regular Polygon*

Ziggy had a date, but Igor, Bob, Chang, and Mohammed were shooting the breeze about their youthful mathematical misadventures.

"Man, did I get in trouble back in seventh grade," chortled Mohammed, "My teacher, whose name I thankfully cannot recall, asked us to calculate the area of a 3-centimeter by 2-centimeter rectangle. I naturally could not believe that the teacher would degrade us with so simplistic a query, so I thought about it a little too deeply. I remembered that the rectangle was the boundary and not the boundary together with its interior, so I gleefully waved my hand and was unfortunate enough to be called on."

"You didn't give him the right answer," said Igor, barely controlling his laughter.

"Oh, but I did! I said the area was zero square centimeters."

"What did you do when you recovered consciousness?" asked Chang.

"I tried to explain my thinking to my mother, my father, the principal, and the school psychologist. The psychologist recommended counseling for my nascent abhorrent behavior. My parents saved a lot of money for themselves by administering corporal punishment instead."

"Child abuse!" accused Bob.

"There were mitigating circumstances," Mohammed continued. "While he was spanking me, my father posed the problem of finding a formula for the area enclosed by a regular polygon. And he said 'enclosed by.' Also, he didn't hit very hard."

"What was the big deal about that?" asked Chang, looking up from the cryptic puzzle at the back of the current issue of *Harper's,* which he had picked up early in the conversation.

"I was only in seventh grade and I was having a bit of a problem thinking about the general polygon, whether or not it was regular. I was standing by my desk because I didn't want my father to see how comfortably I could sit. 'How do you draw one?' I asked him."

"You'd have to draw something open, with dots, but is that the way you want to go? How would you pick up a snake?"

" 'Quick and hard, I'd grab it below the mouth and at the tail,' I told him."

" 'But not in the middle,' he said, fixing me in his intense stare."

Each of the young men had become thoroughly engaged in the problem Mohammed was describing. "He was telling you to look at the extremes!" shouted Igor. "That's great."

Bob looked puzzled. "So, the equilateral triangle is just under the head, but there's no tail."

"Sure there is," Chang said. "The infinite-sided regular polygon. Also known as . . ."

"A circle!" screeched Igor.

"Right," Mohammed resumed. "We'd learned that the area enclosed by a circle was πr^2 and that the area enclosed equilateral triangle is $(s^2\sqrt{3})/4$. Unfortunately, that didn't lead me anywhere. I fooled around with it for a while and something hit me: Circumference was analogous to perimeter. 'Holy moley!' I shouted. The circumference of a circle is $2\pi r$, and its radius is r. Suppose I multiply them and then divide by 2."

Chang was grinning. "Perimeter times radius divided by two. I'll bet you tried the same thing with the equilateral triangle. The perimeter is $3s$, but what did you take for the analogy to the radius?"

"Well," Igor mused, "there's no problem defining the center of an equilateral triangle, so I guess you would want the length of the line segment joining the center to an appropriate point on the boundary."

"A vertex?" asked Bob.

"Hardly!" said Mohammed, with just a tinge of patronization. "How does a radius intersect a circle?"

"Of course," Chang explained, "it intersects the circle orthogonally, so you need the line segment joining the center of the equilateral triangle to the side which meets the side orthogonally."

"And the length of the segment of the altitude joining the center of an equilateral triangle to a side is one third the length of the altitude, or $1/3 \times (s\sqrt{3})/2$. So I multiplied the perimeter by the radius-analogy and divided by two!"

"Getting $(s^2\sqrt{3})/4$," said Bob, with dawning comprehension.

"So I figured, if it worked at the head and it worked at the tail, it might very well work for the whole body. I conjectured that the area enclosed by a regular polygon was always perimeter times radius-analogy divided by two."

"So that's why the apothem is important. It's the radius-analogy in any regular polygon," said Bob. "I think this is the first time I really understood that formula."

Chang smiled. "There is another point that should be considered. Would this procedure still work if we were talking about perimeter instead of about area? Have any of you seen that old chestnut in calculus where a student who had just learned integration attempted to approximate the length of the line segment from the origin to the point (1, 1) by constructing 'staircase approximations'? Each step consisted of a vertical segment with one endpoint on the diagonal segment and a connecting horizontal segment from the top of the vertical segment back to the diagonal. The students then let the length of each vertical and horizontal segment approach zero and claimed that the limit of the sums of vertical and horizontal segments would be the length of the diagonal. Of course, the length of the diagonal is $\sqrt{2}$ and the sum of the lengths of the vertical and horizontal segments is 2, no matter how small you make them."

"Wait a minute," interrupted Mohammed. "Length approximations must be made with *chords!* Both ends of the chord must be on the curve or line segment being approximated. That's not the case in your example."

"Exactly," said Chang. "And if we approximated the arc length—that is, the circumference of a circle with chords—we wouldn't have a problem. Since the length of a side of a regular n-sided polygon is $2r \sin(\pi/n)$ where r is the length of the segment joining its center to a vertex (the central angle formed by two 'radii' and a side has measure $2\pi/n$ so that half that angle has measure π/n, and it follows by right triangle trigonometry that $\sin \pi/n = (s/2)/r = s/2r$). Thus, the perimeter is $n \times 2r \sin(\pi/n)$. Those students who have learned that $\lim [(\sin x)/x] = 1$ as $x \to 0$ would let $x = \pi/n$ and see that as $n \to \infty$, $x \to 0$. Thus as $n \to \infty$, $ns = 2rn \sin(\pi/n) = [2r\pi \sin(\pi/n)]/(\pi/n) \to 2r\pi \times 1 = 2r\pi$, as we would wish. If you know the ideas involved, it all works."

"I wish all of life were like that," said Mohammed wistfully.

② ⑥ ⑦ GEOM

5.2. *The Line $\overleftrightarrow{l}$, the Line Segment $\overline{AB}$ Disjoint From $\overleftrightarrow{l}$, and the Angle of Maximum Measure $< APB$, With P on $\overleftrightarrow{l}$*

Big Ed the bookie had had a profitable day. The favorites all had lost and he had given short odds on the underdogs. He double-locked the metal doors and bolted the reinforced shutters.

"I've done enough arithmetic; time to relax with mathematics." Big Ed opened his well-worn copy of George Polya's *Mathematics and Plausible Reasoning*. He was working on the problem of finding the angle of maximum measure formed by the endpoints of a line segment and a variable point P moving on a nonintersecting line $\overleftrightarrow{l}$ in the same plane. His first conjecture had been that the point P yielding the angle of maximum measure was the projection of the midpoint of $\overline{AB}$ onto $\overleftrightarrow{l}$. His flush of pride was utterly dispelled by a healthy dose of reality.

"It can't be," he censured himself. "Winner and straggler." He sketched two diagrams on his notepad. In one, $\overline{AB}$ was parallel to $\overleftrightarrow{\ell}$; in the second, $\overline{AB}$ was perpendicular to $\overleftrightarrow{\ell}$ without intersecting it. "It works in the parallel case, but it's as bad as it can be in the perpendicular case. If it ain't true for winners and for stragglers, it sure can't be true for all horses.

"Well, as a start, they all got four legs and a body holding them together. How can I tie the line segment to the line in a solid way that's the same for the parallel case and the perpendicular case?"

Ed stared at the page for a very long time. "I sure ain't getting to the winner's circle on this one."

Big Ed's occupation once again served a scholarly purpose. "Circle," he screamed, smashing his often-broken knuckle on the oak table he had inherited from his father, who had met an early unnatural death. Ed drew circles (or slightly distorted approximations thereto) containing the line segments as chords and resting on the line. "I'll bet those are the points," Ed said. He took out his red pencil and drew a second circle, symmetrically placed, in the case of the perpendicular segment. "And not necessarily a oner. Winners don't have to be oners." He thought that his black-sheep Uncle Ted, who had forsaken family tradition to become a mathematics professor, wouldn't be satisfied with talking about a circle with its feet on the ground. Big Ed attempted to mimic Ted's Boston accent: "The angle of maximum measure is formed with vertex at point P where a circle having the line segment as a chord and tangent to the line intersects the line."

"But can you prove it?" Ted chided from within Big Ed's psyche.

"Sure," replied Big Ed, yielding to schizophrenic tendencies. "When the point of tangency is the vertex of the angle, it's measured by half its intercepted arc. For any other point, the angle formed is measured by half the difference between the measure of that same arc and some other arc of the circle, so it's smaller."

"Good lad!" said Ted. "But how can you be sure such a circle always exists?"

Big Ed realized immediately that he should think about a way to construct a circle having a given chord and tangent to a given line in the same plane which did not intersect the chord.

"Easy in the parallel case, so let's look at the straggler." Ed figured that if he could do the construction in the perpendicular case as well as the parallel case, he would be able to work out a construction for the general case.

"I can see," Big Ed continued in his odd discourse, "that if we examine the line containing $\overline{AB}$, it meets $\overleftrightarrow{\ell}$ at point Y. The circle I seek passes through A and B and is tangent to $\overleftrightarrow{\ell}$ at point P to be determined. So $\overline{YBA}$ is a secant of that circle, and $\overline{YP}$ is a tangent." Big Ed remembered his high school plane geometry class, which he loved more than any other class he had taken in school but which had been replaced in recent years by some atrocity containing algebra, geometry, and several other uninteresting topics. "From a point external to a circle in the plane of the circle, a tangent segment is the mean proportional between the entire secant segment and its external subsegment. It's pretty easy to construct a mean proportional using a segment of length $YB + YA$ as the hypotenuse of a right triangle with the vertex of the right angle at a point at distance YB from the endpoint of the hypotenuse. Once I have the length of the mean proportional, I can find the location of P immediately, giving me the angle APB of greatest measure."

"Ed, your language does not conform to modern considerations of linguistic precision," Ted chastised. "You should be careful to say 'the length of the tangent segment . . .' "

"Shut up, Ted! You talk so much that nobody listening to you can conceptualize the content of what you say at length." Ed chortled at his own play on words. "You talk your way; I'll talk mine."

"Your tirade is an ignoble attempt to cloud your own confusion. You are able to construct point P in the extreme cases; you have learned little about what to do in the general case."

"Then, you're a pretty bad teacher, Ted. Didn't you always say that we should look at extreme cases to see if a pattern is established that lets us know about the general case?"

"The key word, Ed, is *if*. In this case, examination of the extreme cases tells you very little about the general case, except for the very beautiful insight that *one* of the two extremes immediately generalizes. That's often the case, Ed. Even when the extremes do not give you a complete insight into the proof or pattern you seek, you have not wasted your time by looking at them. In this case, your insight about the length of the tangent segment and the length of the whole secant and of its external segment was in no way related to perpendicularity. Your proof generalized immediately. The construction for the parallel situation is a special case."

"So, you're saying I should get in the habit of looking at the extreme cases because they so frequently give insight; you're not saying exactly what insight they will give. I guess it varies with the problem." Ed was realizing that his mathematical investigations for the evening were drawing to a close. His long-dead Aunt Agatha had just walked through the wall with a 5-year old version of his intolerably obnoxious cousin Lucy. "Time for some sleep," said Ed, resting his head on his forearm, which itself was resting on his desk, and drifting off to a better world.

⑪

GEOM

GEN

5.3. *A Mile South, a Mile East, and a Mile North*

Calvin threw his stuffed pet on his bed and began tormenting his father in earnest.

"So you got this bear. He walks a mile south of where he woke up, then walks a mile due east, and finally walks a mile north. He finds himself at the exact spot where he woke up. What color is the bear?"

"That's impossible, Calvin! If you walk a mile south and then a mile east and then a mile north, you can't possibly get back to where you started."

"Ha, Ha! on you. Dad," mocked the charming little boy. "You just aren't smart enough." A low guttural sound began to swell from the father's throat. He was rising from his chair with his hands locked in claw position when Calvin's mother left her bizarre supper pot to intercede.

"You stay put!" she ordered her husband, who obeyed instantly. "Calvin, what have I told you about teasing?"

"Oh! Don't tease dumb animals. Sorry, Dad."

"Calvin claims that it's possible for a bear to walk a mile south, a mile east, and then a mile north and be back where it started," said the belabored father in his victim monotone.

"The bear is white, Calvin," his mother said in her own quiet but triumphant tone.

Calvin had become irrelevant. The father's eyeglasses tumbled to the floor. He stepped on them and shattered them as he rose to face his wife.

"You see, dearest, the bear has to be a polar bear, because the bear had to start at the extreme of northernness."

"Ha ha on you, Mom," Calvin interposed in an unlovely way. "That's not the only solution."

"Of course it is, Calvin! Where else could a bear have started than at the North Pole?"

"You forgot to look at the other extreme, Mom."

"Ha ha on you, Calvin," said the father, in a competitive and distinctly nonparental tone. "You can't go south from the South Pole, so you're wrong."

"I didn't say to go south from the South Pole." Calvin's face exuded evil. "I didn't say to start at the South Pole. Somewhere just above the South Pole there's a little latitudinal circle of circumference one mile. Start anywhere on the larger latitudinal circle one mile north of that and look what happens."

"Ha ha on you, Calvin," said his mother, getting into the spirit of this common form of family discussion. "I can think of lots of other places to start. Suppose you started anywhere on the latitudinal circle one mile north of the latitudinal circle with circumference $1/n$, where n is a positive integer. Just look what happens."

Calvin had a perfect rejoinder in a nascent Oedipal triangle. "Gee, Dad. Mom's the smartest in the family and you're the dumbest."

A look of triumph was growing on the father's face. His tone was cool and arrogant. "You are aware, both of you, that there are no bears down there. Your theorizing was interesting, though of course obvious to me. I rejected it because your results were in conflict with reality, and, after all, Calvin did say that a bear was walking. The only valid answer was the one Calvin gave the first time when I chose to let him feel good."

5.4. Clarification and Discussion

The concept of *extreme* is not always well defined. Even in law, a term such as *extreme cruelty* requires some definition. Must extreme cruelty be lethal? If we looked at the region consisting of the interior of a triangle and its boundary, are the extremes the boundary or are they the vertices of the triangle?

When extremes are defined in a problem situation, it is often productive to examine the problem in the extreme cases. The additional hypotheses available because of the extremization often yield results more easily than study of the general case. Sometimes, the special cases at the extremes lead one directly to deep insight into the general case, as was illustrated in the discussion in section 5.1. In 5.2, the extremes yielded considerable insight into the solution of the problem as initially posed but did not yield direct insight into the construction of the point yielding the angle of maximum measure. Nevertheless, one of the extremes led us to general insight on its own. The peregrinating bear problem might easily be made compatible with Mom's proposed solution by eliminating the bear and posing the problem as follows: What is the (complete) set of points on the surface of the earth from which a perfectly designed earth-skimming electronic pigeon could fly one mile south, one mile east, and one mile north and end up at the starting point of the flight? Note that it is productive to think of the extremes, the North Pole and the South Pole, even though only one of those yields an immediate answer. By thinking of the South Pole, we make it easier to visualize the set of circles near the South Pole, which provide an uncountably infinite number of additional solutions.

The advice "Don't go to extremes" is unsuitable for mathematicians.

6. Visual Representation

The quality of many experiences is ineffably enhanced by appropriate pictorial representations. So, with cognitive processes in mathematics, diagrams, tables, charts, graphs, and spreadsheets are not always *needed*, but they always make things more pleasant.

GEN 6.1. *A Mess of Data: New York Summers*

Only two students had signed up for Dr. Berg's geography class during the past decade, but the school had a rare surplus of funds, and both of them had signed up during the current semester. The course was conducted for the two of them.

"The exciting question is whether there has been a statistically significant warming trend over the past century along the Atlantic Coast of the mideastern states."

"Ah cum from Geohgiah," cooed Sweet Caroline.

"That's been something I've been thinking about every night for the past 10 years," said Hornet, adjusting his horn-rimmed glasses with the tape across the bridge.

"Ah declah, ah nevah met a boy lak y'all." Even Caroline recognized that her attraction to Hornet could be taken as certain evidence of insanity, but she had learned long ago, in the wooded suburbs of Atlanta, never to resist an impulse.

"I would collect weather information from government reports, almanacs, and newspapers going back as far as possible for a fixed site such as New York

City, and I'm sure it would be fairly obvious whether or not there would be climactic, I mean, climatic warming trends."

"Could we all do a joint project on this?" said Caroline, forgetting to drawl.

Dr. Berg was a kindly man and saw many advantages in a joint project, not the most of which was there would be half as many papers to grade. "I think that's a splendid idea. Get right on it."

A pattern of sporadic attendance by the class developed almost immediately. On the day their first drafts were due, a 213-page report was deposited on Dr. Berg's desk, collapsing the fragile right-front leg. Years of memoranda spilled to the floor.

"That's a beautiful blue sapphire ring you're wearing, Caroline," Dr. Berg said to distract himself from the rage and frustration threatening to overcome him.

"We all are engaged," Caroline explained.

"We've put a lot of thought into this paper," Hornet said proudly. "I've been thinking about it during daylight hours for 2 months."

That evening, as Dr. Berg began perusing the paper his prize (and only) students had submitted, he noted with marked dismay that 207 pages of the 213 consisted of raw data organized by source. The sentence concluding the report stated that no conclusions were obvious.

"They will be pleased to note that this requires a rewrite," Dr. Berg thought. "I may even be able to teach them something important."

When Hornet and Caroline reached class that Thursday, they were puzzled to see the chalkboard headed "Compare and Contrast." To the left were 3 of their 207 pages of data, held on the magnetic green chalkboard by cardioid magnets that Dr. Berg had purchased at the Mathematical Notion Shoppe. To the right was an artistically designed set of line graphs in each of the colors of the rainbow.

"That would look just beautiful over the fireplace in Daddy's mansion," said Caroline. "What is it?"

Hornet knew. "It's a multiplex linear graphical representation of the data to the left."

"So why do all the graphs go up?" asked Caroline, with the first genuinely puzzled expression that had ever crossed her face.

"Upward tendencies," exploded Hornet, who had seen all the implications of Dr. Berg's astute pictorial representation of his suggested revision of their masterpiece.

"Hmmmm," said Caroline knowingly.

6.2. *The Seven-Game Championship Series: How Likely Is It That There Will Be a Seventh Game?*

③ ④ ⑩ ⑪ ⑬

PROB

GEN

Big Ed had a lot of action on the NBA Eastern Division championship. A Knicks-Bulls series brought out the high rollers like nothing else. The handle would be well over a million dollars if the series went the distance.

Sugar Pete was an eager apprentice. "The series is sure to go six," he asserted.

Big Ed stopped short. "Y'know, that's an interesting mathematics problem. Let's assume these teams are evenly matched. Just how likely is it that the series goes seven games?"

"Well, the probability of a four-game series is clearly 1/2 × 1/2 × 1/2 × 1/2, or 1/16," said Sugar Pete.

"Wrong!" mocked Ed, "Someone's gonna win the opener, and it's just a matter of figuring the odds on their winning the next three, so the odds on there being a four-game series is 1 × 1/2 × 1/2 × 1/2, or 1/8."

"So the fair odds are 7 to 1. We could give 5 to 1 and swing a profit."

Big Ed smiled. "Put it on the board and see who bites."

"So how do you figure the odds on a five-game series?"

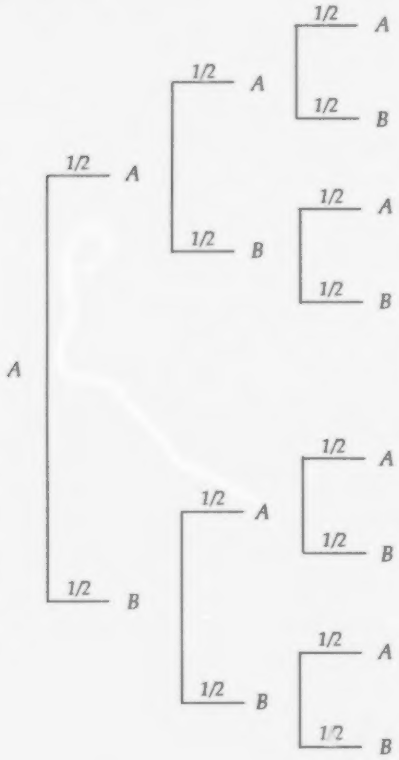
Big Ed eyed his apprentice shrewdly. "You figure it out. Unless you construct your own knowledge of this, you're gonna keep getting it wrong."

"I don't like plain numbers," Sugar Pete whined. "It's much easier with pictures."

"So make a picture," laughed Big Ed. "Look!"

Big Ed started a standard tree diagram with the branches opening to the right. "Let's call the teams *A* and *B* so we don't have to get into a fight over who is going to win the first game. We'll call the winner of first game *A*."

The diagram looked like this:

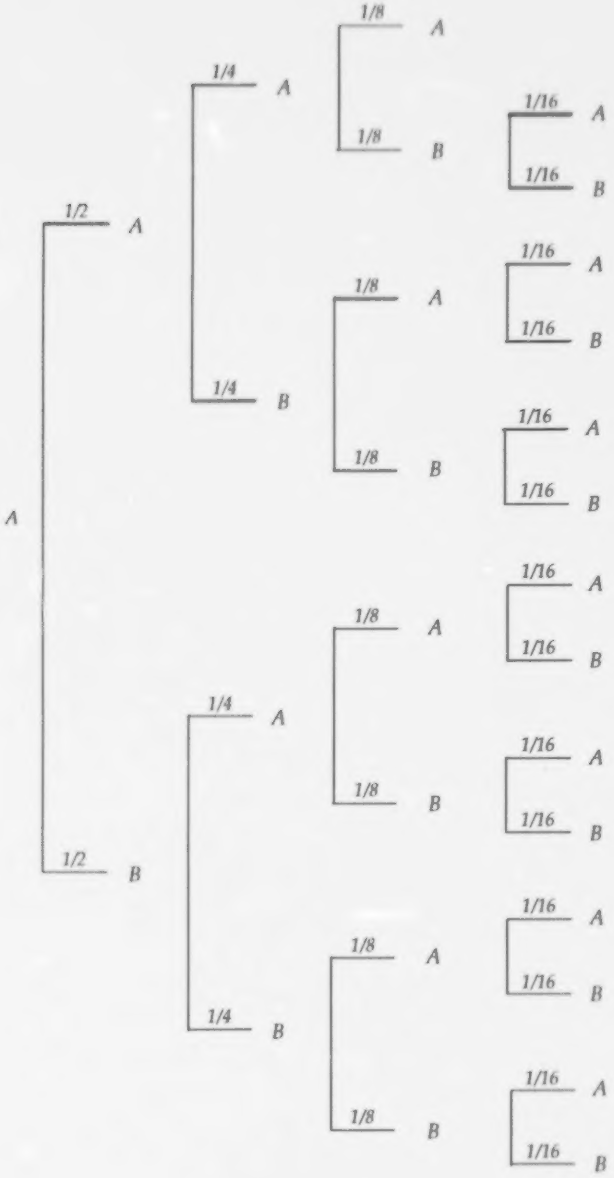


Ed handed the diagram to Sugar Pete. "Fill in the fifth game."

"The top branch is dead," said Sugar Pete. "It's getting complicated."

"Just fill in the live branches."

Pete fiddled around for a while under Big Ed's careful eye. After about 10 minutes, the diagram was filled in through five games.


A
1/2 A
1/2 B
1/4 A
1/4 B
1/4 A
1/4 B
1/8 A
1/8 B
1/8 A
1/8 B
1/8 A
1/8 B
1/8 A
1/8 B
1/16 A
1/16 B
1/16 A
1/16 B
1/16 A
1/16 B
1/16 A
1/16 B
1/16 A
1/16 B
1/16 A
1/16 B
1/16 A
1/16 B

"So what's the probability of a five-game series?" asked Sugar Pete.

"Put a checkmark on all the branches that are in a completed five-game series, and add up the probabilities."

Pete did as he was directed. He found three branches in which *A* won in exactly five games and one branch in which *B* won in exactly five games, so that the probability of the series going exactly five games turned out to be 4/16 or 1/4. Sugar Pete wrote it down.

"The probability of the series going exactly four games is 1/8, and the probability of its going exactly five is 1/4."

"Continue the live branches," said Big Ed, "and figure out the probability of the series going exactly six."

"It's 4/32 that *B* wins in exactly six and 6/32 that *A* wins in exactly six, making 10/32 or 5/16 all together."

"So we can add up 2/16, 4/16, and 5/16 and subtract from one to find out the probability of the series going exactly seven."

"Sugar, you'll never make it as a bookie if you do that. There'll be big bucks riding. Figure out the probability of the seven-game series the way we were going and *check* the result, the way *you* suggested."

Sugar Pete continued the diagram to the seventh column, extending only the live branches. "Let's see. If *B* wins, the branch starts with *A*, ends with *B*, and has five places in between, with three *B*s and two *A*s. Each branch has probability 1/64, so the probability of *B* winning in exactly seven is 10/64. For *A* to win in seven, the branch must start and end with *A*, so we have five spots in between with two *A*s and three *B*s, again coming out to 10/64. All together, it totals 20/64 or 5/16, which checks."

"Bet you hadn't realized that a six-game series is just as likely as a seven-game series."

"Actually, that's obvious," said Sugar Pete. "If the series ain't over in five, it's three to two somebody. Half the time, it will end in six, and half the time it won't. I never thought of it that way before."

"You got some talent. Let's put betting on the length of the series on the board: 2 to 1 on a six-game series and 3 to 2 on a seven. We'll bend the odds to fit the popular misconception. The suckers won't work it out and we'll up our vigorish."

Sugar Pete was thoughtful. "If only they taught them in school to work everything out. If only they taught them to think."

②

GEOM

ARITH

NUMB

6.3. *Adding Consecutive Odd Numbers Beginning With 1*

Nick McBourbak loved second grade. The most wonderful concepts were introduced almost every day. His mom had taught him about odd whole numbers a few years ago and his dad had impressed him with the joy of seeking patterns.

Ms. Clearwhistle was having the collaborative groups explore sums of odd numbers beginning with 1. The kids were moving their counters and charting results.

"1, 4, 9, 16, 25," said Jennifer. "What does it mean?"

"That's a great question," said Nicholas McBourbak with his winning Archimedean smile. "Those are very interesting numbers." He looked down at their chart.

1	1
1 + 3	4
1 + 3 + 5	9
1 + 3 + 5 + 7	16
1 + 3 + 5 + 7 + 9	25
1 + 3 + 5 + 7 + 9 + 11	

"I know the answer before you add them up," Nick stated joyfully.

"You can do that in your head?" asked Mike incredulously.

"That too," said Nick, in a tone which might have gotten him a black eye in another setting. "Actually, I didn't add them at all." Jennifer was impressed. "How'd you get the answer without adding?"

"The pattern, the pattern. The answers are all perfect squares, numbers multiplied by themselves. The next number is 36 and the one after that will be 49."

"And then 64," said Jennifer. "Why is this happening?"

Ms. Clearwhistle could not suppress a smile of triumph. She had never been very good at math, nor her mother before her, but some odd guy had spoken to her methods class and fixed her with a steely gaze. "Never," had he said, a trifle too loudly, "never tell the kids that you were not good in math. Give them good things to think about and the miracles will flow." She was interested in the group's discussion for a profound reason. She had read about this pattern in the national arithmetic journal but had not seen an explanation of why this pattern emerged. "Hot dog!" she thought. "I wonder what the kids can come up with."

Nick had put a yellow counter on the table and surrounded it in a square open pattern with three blue counters. He placed five red counters in a parallel column and a parallel row to form a three-by-three square. Now seven green counters in a new parallel right column and a new parallel bottom row produced a four-by-four square. "Match the column on the right and the row on the bottom and add one in the corner; you get two times something plus one—the next odd number and the next square."

Y B R G
B B R G
R R R G
G G G G

The three children glowed. "I feel as if we did something really good," said Jennifer. "Not just finding a pattern, but explaining it. I wonder if mathematicians ever do this."

"Maybe late at night after they finish adding, multiplying, subtracting, and dividing," said Nick sadly.

6.4. *The Big Picture*

Diagrams, charts, graphs, and figures permit more intensive analysis of data because they permit the purely rational processes involved in higher-order thinking to be combined with direct sensory input. Often, something hidden in its abstract form may blossom in a linear or quadratic curve familiar from previous study. When data form a linear graph, the question, Why are we getting a linear graph? becomes easier to pose and hence possible to investigate. At the highest levels of mathematical research, purely abstract proofs must stand on their own, edifices built from the bricks that are the axioms and held together with the cement that is logic. Often, however, those edifices could not be built without the secret picture relegated to the mathematician's desk drawer after publication of the pure and pristine proof. Yet without that hidden picture or graph or figure or diagram, many proofs would never be. It is the job of mathematics teachers to help students learn how to construct knowledge, and using visual representations is a powerful skill that should be mastered young.

7. Intelligent Guessing and Testing

GAME

7.1. *I'm Thinking of a Number Between 1 and 10,000*

"Dad, I'm thinking of a number between 1 and 10,000. Guess what it is," challenged Calvin.

"The square root of two," said his father, without looking up from *The Wall Street Journal.*

"The who what of which?" Calvin asked.

Now the father looked up. "I take it back. My guess is 2,346."

"WRRONNNGGG!" screamed Calvin. "Guess again."

"One less than one more than the number you're thinking of."

"I'm not thinking," said Calvin, in an unintended admission that made the Flame of Truth on Olympus flare.

"Raw guesses really are quite useless," said the father, kindly. "If I were a stranger and asked you to guess what my favorite song was, what would be the point?"

"I'd get to know you better," said Calvin, regaining his aplomb.

"There's a world of difference between a guess, an intelligent guess, and a studied insight. When there is a good reason for restricting your guess to a relatively small set of possibilities, there is something to talk about and something to be learned!"

"I'll bet your favorite song came out when you were dating Mom," said Calvin. Calvin's father looked amazed. Was it possible that his son had understood what he was talking about when he himself had not?

ALG

7.2. *Solving Equations*

The problem on the blackboard read: "Find a solution of the equation $4(7x + 5) = 9x + 1$."

Sputnik tried to guess: "Let's try 3." He calculated the value of the left-hand member and obtained 104. Hopefully, he now calculated the value of the right-hand member and obtained 28. "Darn," he expostulated. He wasn't too hopeful that his next guess would work out any better. "How can I make a better guess?" he wondered. "Whatever I guess, the left-hand member would end up being 28 times my guess plus 20, and the right-hand member would be 9 times my guess plus 1. If I cut some apples from both sides and then cut some bananas, I guess 19 times my guess would have to equal negative 19. I'm gonna guess negative one," said Sputnik. "I wonder if it'll check out."

How many teachers and students realize that the processes so long taught for solving equations were, in fact, techniques for making a very educated guess, and that the proof that a solution had been found was the verification of that educated guess?

Several years later, Sputnik confronted a new equation in a different class. "Find all real numbers such that the square root of one minus x equals the square root of x minus one."

"Easy as 3.14159," thought Sputnik. He squared both sides. "My guess must have the property that $x - 1$ equals $1 - x$, or twice x equals two, so x is one." He checked it out and was pleased to determine that, in fact, when x had the value one, the equation became zero equals zero.

Unfortunately, there was a second equation to be solved. "The square root of x minus two equals the square root of negative one minus x."

"If you can do one of 'em, you can do 'em all," snickered Sputnik silently. "I'll guess that my solution must have the property that $x - 2$ equals $-1 - x$, or that twice x is one. I guess that x is $1/2$, and I'm tired of always checking, so I'll simply say that as usual it checks out."

He handed in his warm-up exercise, confident of his usual perfect score. When he got his quiz paper back the next day, Mr. Sneed had written snidely in turquoise ink: "What do you suppose is the value of the square root of negative three halves?" For the first time in his uneventful intellectual life, Sputnik began to realize that there was a distinction between an educated guess and a certainty.

② ③ ⑥ ⑪ PROB

7.3. *How Many Heads in Ten Tosses of a Fair Coin?*

Not only was Rip Larkin the best athlete at Franklin Pierce High School, but he was also far and away the most popular boy, and, of course, he knew everything. After Pierce had been defeated in the first four games of their football schedule, Larkin knew that their fortunes would have to improve, because of the Law of Averages.

"I beg your pardon, sir," said Bert Hill, the water boy for the football team. "Things just don't catch up. Suppose you tossed a fair coin 10 times. What would happen?"

"That's obvious!" screamed Rip, "you'd get five heads and five tails in some order."

"Have you got a coin?" asked Bert.

Rip took a quarter from his pocket and handed it to Bert. Bert tossed it 10 times and got a total of 3 heads and 7 tails.

"I guess this wasn't really a fair coin," said Rip apologetically.

Bert had never heard Rip sound apologetic. Buoyed by this new experience, he continued: "The coin may or may not be fair. In fact, a perfectly fair coin might be thrown 10 times, and might come up all heads, although that's not very likely."

Rip was silent. That, too, was a new experience for Bert. "You see, when you toss a fair coin 10 times, there is a non-zero probability for every possible combination of heads and tails. Let me make a chart."

Bert drew a two-column chart diagram and labeled the left-hand column "number of heads in 10 tosses of a fair coin" and labeled the right-hand column "probability of occurrence." He filled in the top row and the bottom (eleventh) row of the chart immediately. Next to 0 he wrote "$1/2^{10} = 1/1024$," and next to 10 he wrote the same equation.

"I'll fill in the 1 row and the 9 row next." He wrote "${}_{10}C_1 \times (1/2)^{10} = 10/2^{10} = 5/512$" on both appropriate lines.

"What does ${}_{10}C_1$ mean?" asked Rip.

"It's the count of the number of length-10 head-tail strings which have a total of 1 head and 9 tails, or one head somewhere in the string of 10. Since there are 10 places you can write the H, ${}_{10}C_1$ equals 10."

"I guess you're going to write the 2-row and 8-row next. I suppose you'll write ${}_{10}C_2 \times (1/2)^{10}$. But it's not clear what ${}_{10}C_2$ is numerically."

"Well, there are 10 places you can write one of the Hs, and for each of them there are 9 places where you can write the other, but you'll end up counting everything twice, because there's no difference between the Hs. Therefore ${}_{10}C_2 = (10 \times 9)/2$, or 45.

"Let's see. I guess the 3 row and the 7 row will be ${}_{10}C_3 \times (1/2)^{10}$ and ${}_{10}C_3$ equals 10 times 9 times 8, all divided by 2."

"Why are you dividing by 2?"

"You said everything is counted twice, so I'm dividing by 2."

"That was when we had two heads. Now we have three. Think of them as a red head, a white head, and a blue head. There are six different ways to write them, so now we must divide not by 2 but by 6."

"I think I'm getting it. So ${}_{10}C_3$ is 10 times 9 times 8, all divided by 6, or 120. Let me try the 4 row and the 6 row myself. You should write '${}_{10}C_4 \times (1/2)^{10}$,' and ${}_{10}C_4$ would be 10 times 9 times 8 times 7 divided by the number of ways you could arrange a red head, a white head, a blue head, and a blonde." Rip reflected on this. "I should say yellow head. Anyway, I can put any of the four colors first, and any of the remaining three second, and any of the remaining two third, so I guess there are 4 times 3 times 2 times 1, or 24 ways to do it, so ${}_{10}C_4$ must be 10 times 9 times 8 times 7 over 24, or 210."

"Which brings us to the middle."

"Wow!," cried Rip. "I see the pass all the way. On the 5 row, we must write '${}_{10}C_5 \times (1/2)^{10} = (10 \times 9 \times 8 \times 7 \times 6)/(5 \times 4 \times 3 \times 2 \times 1) \times (1/2)^{10} = 252 \times (1/2)^{10}$.' " "Shouldn't we not have simplified?" asked Rip. "It would be simpler to check our calculation if all of the denominators were 1,024."

"How would you check?" asked Bert.

"If you left all the denominators as 1,024, I'd add up all the numerators and see if the sum was also 1,024, because something had to happen, and we've taken care of all the cases."

"No wonder you're all-state left end," said Bert, grinning from ear to ear.

"Now that we've got the chart, we can really find out what's going on."

Number of Heads in 10 Tosses of a Fair Coin	*Probability of Occurrence*
0	1/1,024
1	10/1,024
2	45/1,024
3	120/1,024
4	210/1,024
5	252/1,024
6	210/1,024
7	120/1,024
8	45/1,024
9	10/1,024
10	1/1,024

"First, let's check the sum of the numerators. 1 + 10 + 45 + 120 + 210 + 252 + 210 + 120 + 45 + 10 + 1 = 2(1 + 10 + 45 + 120 + 210) + 252 = 2(386) + 252 = 772 + 252 = 1,024. Bingo."

"Not much to choose from among 4, 5, or 6 heads," said Rip, with some amazement. "If someone asked me how many heads they'd get if they tossed a fair coin 10 times, my educated guess would be 4, 5, or 6, but it's no certainty even then. By the way, something's been bothering me since you started the chart. Why didn't you write ${}_{10}C_6$ in the 6-row instead of ${}_{10}C_4$? Were you counting the 10-chains with 4 tails instead of the 10-chains with 6 heads, and simply recognizing that you were counting the same thing?"

"It's really a shame that you dropped math after your sophomore year," said Bert, with great sincerity. "You've got quite a flair."

"I don't have a flare," said Rip. "I'm a left end, not a quarterback."

⑥ ⑬ NUMB

7.4. The Chinese Remainder Algorithm

Ace was the tennis pro at West Shore Tennis Club. He had eight classes of students, each with a different number of persons. Strangely, the sizes of the classes formed a consecutive sequence of natural numbers starting with two and ending with nine. He was paid by each student according to the number of students in that student's class, so that he got the same amount of money for each class, but that's another problem. He wanted to order eight boxes of tennis balls, one for each class, and in each case to divide up the number of balls in the box equally among the students in the class.

"I wonder how many balls I should order in a box so that I can divide them equally in each case." He multiplied out 2 × 3 × 4 × 5 × 6 × 7 × 8 × 9. "That's a lot of balls," he said. "I wonder if I could do with fewer." He decided to ask one of his students, a Mr. Sperling, who was some sort of a math teacher.

"You really only have to worry about the primes," said Mr. Sperling. "If the number of balls in the box is evenly divisible by 2, 3, 5, and 7, you only need worry about 4, 8, and 9, and putting in two extra twos will take care of 4 and 8, and one extra three will take care of 9."

"Huh?" said Ace.

"So I guess you want to order 2 × 2 × 2 × 3 × 3 × 5 × 7 balls in each box," said Sperling.

"How many is that?" asked Ace.

"How should I know?" said Mr. Sperling. "I'm a math teacher, not an accountant. But," he added, "it's over 2,500 balls, and I don't think it's practical. Why don't you order a smaller number, give the same number to all the students in each class, and keep the extras for your special students who take private lessons? If you ordered boxes of 50, it would work out quite well. You give 25 to each of the students in your class of two, 16 to each of the students in your class of three, 12 to each of the students in your class of four, 10 to each of the students in your class of five, 8 to each of the students in your class of six, 7 to each of the students in your class of seven, 6 to each of the students in your class of eight, and 5 to each of the students in your class of nine, and you'll have 14 balls left over for your private lessons."

"I hate having different numbers of balls left over from each class. Couldn't we even it out? Say, what if we had exactly one to be left over from each box after the partition?"

"Are you asking, 'What is the smallest number that yields a remainder of one when divided by each of the natural numbers from two through nine?' "

"I guess so."

"I'm travelling to Beijing next month; I'll find out for you."

"I don't get it," said Ace.

Mr. Sperling laughed at his own wonderful joke. "We mathematicians call that kind of question an application of the Chinese Remainder Theorem."

Mr. Sperling had, in fact, never understood the Chinese Remainder Theorem very well, but he loved the name. One of his former students, who was always very good in math, was coming up for open school night, and he thought he could get the answer with very little trouble.

"Good to see you again, Polly Sue. You know, Nick is so good that he's being moved to a specialized school next year. He developed some result about eigenvectors that nobody in the school understands, and he's always right about everything, and I liked you a lot better than I like him. Oh, by the way, what's the smallest natural number that yields a remainder of one when divided by each natural number from two through nine?"

"I'm probably the first parent in recorded history to be asked about the Chinese Remainder Theorem on Open School Night," said Polly Sue, smoothing her hair. She went up to the blackboard and wrote: "$n \equiv 1 \pmod 2$, $n \equiv 1 \pmod 3$, $n \equiv 1 \pmod 4$, $n \equiv 1 \pmod 5$, $n \equiv 1 \pmod 6$, $n \equiv 1 \pmod 7$, $n \equiv 1 \pmod 8$, $n \equiv 1 \pmod 9$." She then opened her handbag and removed a carefully wrapped stick of yellow chalk. "You still don't use colored chalk properly, and it's been many years since I first told you about it, way back when Jerry was in your class."

Polly Sue had not really been Mr. Sperling's student, but he liked to take credit for both her and Jerry, and no one had been around long enough to challenge the inaccuracy, except, of course for Polly Sue and Jerry themselves. He watched as Polly Sue neatly crossed out the irrelevant equations on the blackboard with the colored chalk. What was left was: "$n \equiv 1 \pmod 5$, $n \equiv 1 \pmod 7$, $n \equiv 1 \pmod 8$, $n \equiv 1 \pmod 9$."

"Why are you crossing the others out?"

"Because they're irrelevant," replied Polly Sue, finally cognizant that she was answering his questions as well as he had answered Jerry's so many years before. "Think about it for a minute and it'll be obvious."

"I see," lied Mr. Sperling. "But how are you going to solve even those four simultaneous congruences?"

"Using the Chinese Remainder Algorithm, of course. The key number is 5 times 7 times 8 times 9, or 2,520. So we know that number is evenly divisible by 5, 7, 8, and 9, and hence also by 2, 3, 4, and 6. One more than it solves the simultaneous congruences."

"But how do we know that 2,521 is the least solution to the simultaneous congruences?"

"If there were a smaller solution, then one less than that solution would be evenly divisible by 5, 7, 8, and 9. Those are pairwise relatively prime numbers. Without going into a formal proof, surely you can see that the least natural number evenly divisible by both 5 and 7 must be 35. It's a mere baby step from there to getting that the least natural number evenly divisible by 5, 7, and 8 is 280, and when you throw in the 9, there you go. Here's the real problem. Instead of making all the remainders equal, suppose I asked you the following question: What is the least natural number n such that n is congruent to $a - 1$ (mod a) for $a = 2, 3, 4, 5, 6, 7, 8, 9$?"

"Could you try to explain your question more clearly? I'm not sure I understand you completely."

"Let me confuse you a little more. Actually, the problem I posed is not the general case at all, because I'm asking you to find n such that n is congruent to -1 (mod a) for $a = 2, 3, 4, 5, 6, 7, 8, 9$, and the only real complexity is that the moduli are not all pairwise relatively prime. For example, $n \equiv -1$ (mod 9) forces $n \equiv -1$ (mod 3), and $n \equiv -1$ (mod 8) forces $n \equiv -1$ (mod 4) and $n \equiv -1$ (mod 2). The only relevant moduli on the list are 9, 8, 7, and 5. I'm sure even you can see that $n \equiv -1$ (mod 2) and $n \equiv -1$ (mod 3) forces $n \equiv -1$ (mod 6)."

"I'm sure."

"So, the problem is easily solved by multiplying 9, 8, 7, and 5 and subtracting one. This is still not the general case."

"Do give me an instance of the general case."

Polly Sue was deeply amused, but she did not betray it in her voice. "I'll make it relatively easy. Find the smallest natural number n such that $n \equiv 1$ (mod 5), $n \equiv 3$ (mod 7), and $n \equiv 4$ (mod 11)."

Mr. Sperling truly astounded Polly Sue with his next remark, until she analyzed that it could have been made without deep comprehension. "Isn't the key number now of the form 5 times 7 times 11?"

"And just what will you do with that key number?"

At this point, Mr. Dugan, another parent, leaped out of his seat, screaming uncontrollably, "When is it my turn to talk about my son, Stubby? You've been talking about that Chinese kid of yours long enough."

Four days later, Mr. Sperling received a neat letter from Polly Sue, which read as follows:

Dear Mr. Sperling,

I'm presenting you with this explanation just so that you can check your own investigations against it.

Since 5, 7, and 11 are pairwise relatively prime, there is no danger of contradictory information. There are four key numbers in the analysis of the problem I posed to you on Open School Night. They are 5 × 7, 5 × 11, 7 × 11, and 5 × 7 × 11.

5 × 7 is important because it is congruent to 0 (mod 5) and congruent to 0 (mod 7), but not congruent to 0 (mod 11). It actually happens to be congruent to 2 (mod 11). We want n to be congruent to 4 (mod 11), so multiplying 7 × 5 by two will give us a number that is congruent to 4 (mod 11) and congruent to 0 (mod 5 and mod 7).

Similarly, we will be concerned with the congruence of 5 × 11 (mod 7). This product is congruent to 0 (mod 5 and mod 11) and is congruent to 6 (mod 7), and we are looking for an n that is congruent to 3 (mod 7). One way to put the appropriate question is to ask, What multiple of 6 is congruent to 3 (mod 7)? You would obtain the correct answer by trial and error and find that 4 times 6 fills the bill. A better way to think about it would have been to note that 5 × 11 is congruent to –1 (mod 7), and hence, since –4 is congruent to 3 (mod 7), 4 times –1 is congruent to 3 in this case.

4 × 5 × 11 will be congruent to 3 (mod 7) and congruent to 0 (mod 5 and mod 11).

Finally, consider 7 × 11. 7 × 11 is congruent to 0 (mod 7 and mod 11) and is congruent to 2 (mod 5). Since we want a number that is congruent to 1 (mod 5), we ask what multiple of two is congruent to 1 (mod 5), and clearly 3 times 2 fills the bill. Thus, 3 × 7 × 11 is congruent to 1 (mod 5) and congruent to 0 (mod 7 and mod 11).

The number $n \equiv 2 \times 5 \times 7$ plus $4 \times 5 \times 11$ plus $3 \times 7 \times 11$ clearly satisfies the three congruences $n \equiv 1 \pmod 5$, $n \equiv 3 \pmod 7$, $n \equiv 4 \pmod{11}$, because each term establishes one of those three congruences and the remaining two terms have no impact with respect to the modulus considered in the original term.

The only worry remaining is whether n is the least such positive integer. In this case, $n = 70 + 220 + 231$ or 521 satisfies the three congruences but is not the least positive integer that does so.

The final key number comes into play. Multiples of 5 × 7 × 11 are all congruent to 0 (moduli 5, 7, and 11). Thus, adding or subtracting any multiple of 385 will not affect the three congruences. We subtract 385 from 521 and get 136, which is the least positive integer satisfying the three congruences.

With suitable affection,
Polly Sue McBourbak
Jerry and Nick

7.5. *Second Guessing*

The discussion of the Chinese Remainder Theorem raises an important issue. Algorithms are the antithesis of guesses; why place a discussion of the Chinese Remainder Theorem in a section on educated guessing and checking?

I believe that much knowledge arises from comparison and contrast, and the nature of the problems posed spans situations in which educated guesses were possible and situations where only certainty would suffice. When all the congruences were the same and only the moduli differed, an educated guess that the solution was near the least common multiple of the moduli was on the money. When the congruences were different, as in the final example, the algorithm seems to be the only feasible alternative to a lengthy trial-and-error process.

Thus, a new question can be posed by teachers and students alike: In what types of mathematical problem situations are educated guesses appropriate?

I urge the reader to think deeply about this. The answer has profound pedagogical significance; the decision as to whether insight is accelerated through guessing is not one for which guessing is appropriate.

Think about Calvin and the number from 1 to 10,000, and about the solution of equations that may or may not have "extraneous" solutions. Think also about the place of estimation in the teaching of arithmetic. The processes whereby guesses become educated guesses and then conjectures are at the heart of the sophistication that we call *mathematical intuition.*

8. Determining Necessary or Sufficient Conditions

⑬

8.1. *Today Is Sunday*

LOGIC

GEN

"What day of the week is it, Sweetcardioid?" Jerry yelled out in his sleepiest voice.

"Yesterday was Saturday, dear Fiber Bundle," replied Polly Sue from the dining room, where Nick was serving her breakfast.

"But if tomorrow is not Monday," shouted Nick, "what can we deduce?"

"That's easy," said Jerry, hugging the best mathematician in the house." When your mother said that yesterday was Saturday, it must have been a few seconds before midnight on Sunday, so that when you spoke, tomorrow had become Tuesday."

"Aha!" said Polly Sue. "When I said, 'Yesterday was Saturday,' how did you know whether I was making a statement of fact or creating a premise for an argument?"

"Right!" Nick chimed in. "It is important to recognize that when information is *given* in a mathematical argument, as you so carefully explained to me, there is no presumption that the given information has any relevance to truth. Mathematical arguments must hold whether or not the premises are true. The sentence, 'If triangle *ABC* is isosceles, then it has at least two congruent angles' is true for right triangle *ABC* with sides measuring 3, 4, 5."

"True," said Jerry, "but remember, Nick: Not only do your mother and I never lie to each other, but we try to the best of our abilities not to be deceptive while telling the truth."

"How did you guys ever stay together long enough to have me?" asked Nick, with genuine admiration for two parents who probably would have been saints in another era. "Truth telling puts an awful strain even on my friendships in school. When I get old enough, we'll really have to talk about this."

"Not until you're at least 25," said Polly Sue, very close to laughter.

"Or just before your wedding," said Jerry, in his most paternal tone.

"The real facts of life," said Nick.

"Anyway, if your mother had meant 'Yesterday was Saturday' as a premise, I would have expected her to say, 'Suppose yesterday were Saturday.' When she actually said, 'Yesterday was Saturday,' I took that as a statement."

"But Fiber dear, yesterday is a pronoun. No sentence including a pronoun can be a statement in written exposition, and we always swore that we would speak to each other as if our words were being written for posterity. Had I said that July 25, 1997, would fall on a Saturday, that would have been a statement."

Nick was puzzled. "What if my teacher asks me, 'What is your name?' Should I say that 'you' is a pronoun?"

"Of course not," cautioned Jerry. "What teacher speaks with words that should be written for posterity? Furthermore, no question can be a statement. Finally, neither your mother nor I want to go to school again because of your literal interpretations of instructions and questions."

"In common speech, 'Yesterday was Saturday' would be viewed as a statement. Formal mathematics is a written discipline, but it is quite proper to apply logic in spoken arguments."

"Even at my tender age," said Nick, "I have learned the lesson that it may be quite proper to apply logic to spoken arguments, but it does place one at physical risk."

"One thing is certain," said Jerry, attempting to terminate the now-out-of-hand discussion. "If we accept the cyclical sequence of days of the week in their traditional order as axioms, then it follows that if yesterday were Saturday, then today is Sunday."

"Amen," said Polly Sue.

8.2. The Third Easy Month This Year

②
⑬
NUMB

As regional manager of Syndicated Software, Mohammed Hobbes found that memory feats convinced many clients of his special brand of intelligence. He particularly liked the fact that he could project dates and accurately determine the day of the week on which they would fall. For that reason, he was particularly fond of the easy months; namely, those that began on Sunday. "If you're so good," challenged Knute Rockwell, president of Hardware Etc., "tell me the month in which I was born. It was the third easy month in a leap year."

"Before or after Independence Day?" said Mohammed, triumphantly.

"You're not going to get me to tell you which half of the year my birth month is in."

"I wasn't asking about halves of the year. I was asking whether you were born before the Fourth or after the Fourth. I knew immediately that you were born in July, and that you had not been born on July fourth."

"OK, Hobbes! This you are going to explain."

"In every ordinary year, there are four sets of months that have the property that the months in the set must start on the same day of the week. They are {January, October}, {February, March, November}, {April, July}, and {September, December}. In leap years, three of those sets are changed. The leap year sets are {January, April, July}, {February, August}, {March, November}, and {September, December}. This can be seen through an analysis of the number of days in a consecutive set of months added modulo seven."

"What's a modulo?"

"Arithmetic modulo seven is an analog of arithmetic on a clock. Imagine the standard Venusian clock, with hours marked '0', '1', '2', '3', '4', '5', '6.' "

"Is that really the Venusian clock, Mohammed?"

"I ought to know. Consider the ordered triple (January, February, March). On our Venusian clock, 31 equals 3, 29 equals 1, and 31 equals 3, and 3 plus 1 plus 3 equals 0."

"Why did you say 31 twice?"

"For January and March."

"Oh! And February gave you 29 because it's a leap year."

"Right. Since we got a 0 sum, January and April must start on the same day of the week. Now, in any year, let us do the same arithmetic for the ordered triple (April, May, June). Here, 30 equals 2, 31 equals 3, 30 equals 2, and 2 plus 3 plus 2 equals 0, so April and July always start on the same day of the week in a given calendar year."

"But how did you know that I wasn't born on July fourth?"

"I've seen your stationery, Knute. If you had been born on the Fourth of July, you'd have had an American Flag and the logo Yankee Doodle Dandy somewhere on the stationery."

"You bet you. By the way, Mohammed, get the papers together for that $7 million contract to provide your word processing package as standard equipment on the Hardware Etc. PC."

"It's all ready," said Mohammed, pulling a sheaf of papers from his upper-left desk drawer. "Sign here."

8.3. I Am a Parallelogram

GEOM

"Who am I?" asked Hornet, as he gave his report to the students of his Mathophobe I elective. "I am a plane polygon."

"He got that right," said Sweet Caroline, who methodically enrolled in every course that Hornet took.

"I have clearly defined pairs of opposite sides, and the sides in each pair have equal length. I have two pairs of opposite angles, and the angle in each pair has equal measure. My diagonals meet in a single point which bisects each of them. I have at least one obtuse interior angle."

"What self-awareness," mumbled Sweet Caroline.

"Each pair of adjacent interior angles is supplementary. No two adjacent sides of mine are equal in length. Again I ask, 'What am I?'"

"An arrogant male chauvinist pig," thought Sweet Caroline, lovingly.

"A multiredundant parallelogram," said Oscar, about whom the less said the better.

The Master looked up from his desk. "You raise an interesting point there, Hornet. Exactly one of your descriptions was necessary and sufficient for determination of your identity as a parallelogram. Note, for example, that in a regular hexagon all the diagonals meet at a point that bisects each of them, there are clearly defined pairs of opposite sides that are equal in length, and the opposite angles are equal in measure. There are multitudinous polygons in which no pair of adjacent sides are equal in length and an uncountable number of polygons that contain at least one obtuse angle. Of all the conditions you mentioned, only the statement that you are a polygon with each pair of adjacent angles supplementary forces you to be what you profess to be. Naturally, you added something by making yourself both irregular and unrectified, as well as nonrhomboid.

"You raise a fascinating issue about the nature of optimal definitions in mathematics. Should a definition be minimal or should it be maximal? Suppose we put down a large set of conditions on quadrilaterals and proved that any one of them implied each of the others. As an example, consider the following:

Quadrilaterals with

- two pairs of sides equal in length
- two pairs of parallel sides
- one pair of sides both parallel and equal in length
- diagonals that bisect each other
- each pair of opposite angles having equal measures
- each pair of adjacent angles being supplementary

"Suppose we prove that, for quadrilaterals, each of the given conditions is sufficient to prove each of the other conditions. We then say, 'A parallelogram is a quadrilateral with one of, and hence all of, the above conditions.'

"Consider the pedagogical implications. For you math phobic postadolescents, there would be no hidden information. You would have a gestalt of the quintessence of parallelogramism. The mathematics would be in the extensive proof prior to the definition, which used our illustrious discipline to establish the necessity and sufficiency of each listed condition.

"Contrast this with the traditional approach through which each of you has long suffered. Minimal definition hid the full meaning of the word being defined rather than elucidating it.

"Even the commitment to a minimal definition is hypocritical. What is the minimal definition of a regular polygon?"

"That's easy," said the indescribable one (Oscar). "A regular polygon is a polygon all of whose sides and all of whose angles are equal in measure.

"I thought a neater definition was the one a substitute teacher gave us late one Friday. He said, 'A regular polygon is a polygon with a well-defined center such that for every line through the center there exists a one-to-one correspondence between the points on the polygon on one side of the line and those on the other, such that the perpendiculars to the line from each pair of corresponding points are equal in length.' "

"That doesn't work for polygons with an odd number of sides, but it is elegant for those polygons with an even number of sides."

"In any event," said the Master, "the first definition was not minimal and the second was not general. A minimal definition for a regular polygon might be as follows: a regular n-gon is an n-sided polygon in which all of the angles are equal in measure and at least $n - 1$ sides are equal in length. I challenge you to prove that such a polygon is regular, but I assure you that it is. Alternatively, consider the following: a regular n-gon is an n-sided polygon with n congruent sides and at least $n - 1$ congruent angles.

"Neither of the last two definitions I gave to you is psychologically sound. Fewer of you would be math phobic if definitions were formulated not through minimalist criteria but through their psychological soundness. Be children again in high school and think which of the following definitions of a square would be less likely to produce mathphobes?

1. A square is a rhombus with at least one right angle.
2. A square is a rhombus with four right angles.

"The point that I make is this: If two definitions are each mathematically sound, why not select the one that gives a greater instantaneous gestalt to the greatest number of people? Redundancy is superior to obscurity."

8.4. *Another Look at Solving Equations*

ALG

In section 7.2, we raised the point that the standard technique for solving equations consists of making deductions from the equation to be solved that yield necessary conditions for the solution—that is, that yields a set which includes all numbers that are candidates for solutions of the equation. The check is the sufficiency proof, which weeds out the elements of the candidate-set that are unqualified for the office of solution.

Each time a student engages in a standard equation-solving procedure, the student in his deduction and check is giving a proof of necessity and sufficiency.

8.5. *Locus*

⑥

GEOM

"The locus of points satisfying a condition is nothing more than the set of points in a prespecified space that satisfy that condition. Consider acute angle *ABC*. Find the locus of points equidistant from the sides of that angle."

Nick had been working on this problem for 15 minutes when Ms. Gale walked over to his collaborating foursome. "What's going on?" Dorothy Gale inquired. "Everyone else has finished."

Nick looked up. His teacher had come from the heartland of America, but she seemed to lack heart. Nick felt feisty and decided that he wouldn't mind another visit with the principal.

"They may have finished the problem in Kansas, but they're still working in Massachusetts."

On the drive home from the latest in the series of principal's disciplinary conferences, Polly Sue was acutely aware of her son's impending adolescence and was uncharacteristically taciturn.

"Mom, isn't there anywhere where correct mathematics is more important than political sensitivity?"

Polly Sue decided to take a purely academic approach to a delicate question she preferred not to answer. "What problem were you working on?"

"You know Miss Gale?"

"She wasn't in the school when your father and I attended. She had some special history that we talked about some years back, but it escapes me."

"Well, she asked us to find the locus of points that were equidistant from the sides of an angle. I sharpened her question by limiting my inquiry to the plane of the angle. Nevertheless, she thought that the answer was the line containing the angle bisector of the angle, and most of the kids in the class were astonished to consider that the half-ray exterior to the angle was part of the locus. It didn't take me long to see that the problem was considerably more complex than Ms. Gale had intended, and I actually became quite engrossed in my investigation. I guess I got angry when she interrupted me, and I made a disparaging remark about Kansas. In retrospect, and *in toto,* I was off-base."

Polly Sue had not been fully attentive, but for some reason, something in her son's answer struck a chord. "What did you say about Toto?"

"I didn't say anything about Toto. Who's Toto?"

Polly Sue was in that state of proximal development when she could not put the fingers of her mind on the substance of her recollection.

"Anyway, Mom. I realized that to find the necessary and sufficient condition for a point in the plane of angle ABC to be equidistant from the sides of the angle, it would be necessary to define the distance from a point to a ray. Most of the kids assumed that the distance from a point to a ray was the distance from the point to the line containing the ray, but I recognized immediately that that was not right. If we drew a line perpendicular to the ray from its endpoint, and took any point P in the half-plane not containing the ray, the distance from P to the ray would be the length of the line segment joining P to the endpoint of the ray. I drew rays perpendicular to $\overline{AB}$ at B and $\overline{CB}$ at B, each of which formed an obtuse angle with the ray of angle ABC to which it was not perpendicular. Let's call the two rays $\overrightarrow{BM}$ and $\overrightarrow{BN}$. I realized about π seconds before Miss Gale interrupted me that not only was the ray bisecting angle ABC part of the locus, but so was every point P in the interior or on the boundary of angle MBN, and I can prove the necessity and sufficiency of being an element of the union of that angle, its interior and the bisecting ray for assuring that a point is equidistant from the sides of angle ABC. I deserve a principal's citation for this insight."

"You got one," said Polly Sue ruefully.

"So tell me, Mom, why does Miss Gale so often wear red shoes?"

8.6. A Summary Perusal of Necessity, Sufficiency, Truth, and Validity

Investigation of necessary and sufficient conditions for a prespecified set of circumstances is a part of many disciplines but is central to mathematical thinking. Virtually every proof in mathematics is either an existence proof, a uniqueness (or nonunique delimitation) proof, or a proof of either necessity or sufficiency of one condition for another.

It is interesting to speculate upon whether mathematical investigations lead to increased understanding of the real world per se. A valid mathematical proof is one in which each step follows from prior steps, premises, axioms, definitions, or theorems of the system in which the proof occurs, using the established precepts of two-valued logic. When a valid proof has been given, we are in no way claiming the truth of any sentence in the proof. In fact, the axioms of a mathematical system are themselves neither true nor false, because they contain undefined terms. When an interpretation is given to each undefined term in a mathematical system, and it is found that under that interpretation, each of the axioms is true, the network of logically connected sentences in the mathematical system sheds a laser beam on new knowledge and yields truth in unexplored areas that were previously unilluminated. So long as interpretation makes the premises of a mathematical argument true, then we are assured that all the consequents of the argument are also true.

Mathematics may validly be pursued in a black box, but truth requires peepholes into the real world, if there is one on which a community of inquiry can agree. We can argue about the appropriate age at which students should be introduced to axiomatic systems; it is certain that when youngsters first comprehend the beauty of such systems they will begin to see issues of necessity, sufficiency, truth, and validity with new eyes.

9. Sequencing

9.1. *The Chicken or the Egg?*

"The chicken had to come first," said Alice.

"No, it didn't, Miss Liddell. There had to be an egg from which the first chicken hatched."

"But God made the first chicken," said Alice.

"No, God made the first egg."

"So who sat on the first egg?" asked Alice

"The argument is truly pointless," said Herr März. "We can certainly agree that whenever a_n equals chicken, a_{n-1} and a_{n+1} each equals egg. If we let the domain of this sequence be the integers instead of the natural numbers, we avoid the issue of first causes."

"But then, whose head would we chop off?"

9.2. *A Logic Puzzle: Determine the Order*

GEN

Four events are given, and the children are asked to put them in proper time sequence:

1. Jerry wakes up in the morning.
2. Jerry eats supper.
3. Jerry goes to school.
4. Jerry goes to his afternoon job.

Lucy had no problem. Clearly, the correct order was 1, 3, 4, 2. It could be nothing else.

"I don't know," said Franklin." Why couldn't it be 4, 2, 3, 1? My brother Jerry goes to night school."

Patty lifted her head from the desk. "Sequences don't have to begin with the morning anyway; I never really wake up until three o'clock in the afternoon."

"As always," barked the strange boy with the long ears who sat in the back of the class, "your sequencing problems are ambiguous."

Nobody got angry, because nobody understood him. As always.

① ② ④ ⑪ ARITH GAME

9.3. *The Game of 31*

Big Ed had made a great deal of money from the game of 31 over the many years when he played it with the unwary. The rules were simple enough so that his victims came to believe that in just one more play of the game they would win. They rarely did.

Thirty-one is a two-player game with alternating turns. The initial layout consists of cards marked with the numerals 1 to 6 with exactly four of each; all the cards are sorted and placed face up on the table. The victim has the choice of picking first or second. Each player in turn picks a card, places it face down on the table, and adds the number indicated by the card to the current total of previously selected numbers (which is zero at the start of each new game) with the stipulations that when a particular domination is exhausted, that number is not available and that under no conditions may a player turn a card making the total greater than 31.

Ed liked to use the aces through sixes from a standard deck of cards. "Do you want to go first or second?" he asked.

"You go first," said The Victim.

Ed never used the correct strategy at the start. He turned a 6. The Victim thought (about nothing in particular) for a few seconds and turned another 6, saying, "That makes 12."

"So it does," confirmed Big Ed. He now turned a 5 and said "17."

Again The Victim thought. He turned the third 6 and said, "23."

Ed turned an ace in a bored gesture. "24."

For the first time The Victim actually thought about the situation. He realized to his mortification that the $10 wager had been lost. Big Ed would be able to reach 31 exactly to win the game. "I guess 24 wins."

Ed began to become really interested in the situation. The best suckers were always the ones with a little bit of knowledge.

"So you get to 24," challenged the bookie-turned-shark.

The victim knew a *little* mathematics. He began to analyze the significance of 7 in the game: 7 was the least number unmakeable in a single turn. Thirty-one minus 7 was 24. "That is the start of a backward sequence of winning moves," he thought triumphantly. "The sequence is (31, 24, 17, 10, 3)," he continued silently. "I've got him."

"Let's up the wager a little," said The Victim ripely. "Can we make it a hundred?"

"Go slow," said Ed in a feigned cautionary tone guaranteed to reel in his fish.

"Make it a hundred. I want to go first."

"Don't say I didn't warn you," said Ed, thinking about tomorrow.

The Victim turned the first of the four 3s and said, "3."

Ed turned a 4. "7."

The Victim turned the second of the four 3s. "10."

Ed turned a 4. "14."

The Victim began to feel uncomfortable. He turned the third of the 3s, saying "17."

Ed turned the next-to-last 4 and said, "21."

The Victim turned the very last 3 and said, "24."

Ed turned the last 4. "28."

No 3s remained. If The Victim turned a 1, Ed would win by turning a 2; if The Victim turned a 2 Ed would win by turning a 1. The hundred dollars had been lost.

"Tell you what," said Ed. "I'll give you a week to think about the game. We'll meet here next week and play one game for $250. What do you say?"

"You're on!"

The Victim had a plan. He was going to see a client of his named Jerry McBourbak that very evening to discuss a life insurance policy, and that McBourbak was a math whiz. He would ask McBourbak for help.

After he had wrapped up a nice commission, The Victim broached the game to not one but three very interested pairs of ears. He recounted the events of the day with great accuracy.

"He was able to force you to deplete the 3s," summarized Nick.

"Right," said The Victim, still addressing Jerry rather than the two better mathematicians in the room.

"It is apparent that going first and playing 3, 4 or 6 loses," said Jerry. "We see the depletion strategy if you take a 3; the other two moves permit Ed to use his 10, 17, 24, 31 sequence."

"The only time 24 is a winning move is if there are one or more of each denomination left," said Polly Sue.

Nick smiled. "Seventeen wins if there are two or more of each denomination left, 10 wins if there are three or more of each denomination, and starting with 3 would win if there were four or more of each denomination left, which is impossible in Ed's little game."

"So should I ask to go second?"

"Not at all," said Nick as Polly Sue and Jerry beamed with pride. "Turn the depletion strategy against Ed. Start with 5. He'll have to take the second 5 to make 10 or let you get into winning sequence with enough of each denomination left."

"I get it," shouted The Victim erroneously. "Then I take the third 5 and make 15."

"No, no!" said Nick. "Ed will take a 4 making 19, forcing you to take the last 5 to make 24. He will then win by turning a 2 to make 26 and you'll have no 5s left."

"So starting with 5 doesn't work?"

"Sure it works. After Ed takes the second 5 to make 10, you turn a 2 to make 12. That forces Ed to take the third 5 to make 17. Now you turn a 2 to make 19 yourself. He has to take the last 5 to make 24 and you now turn a 2 to make 26 defeating him."

"A word of caution," said Polly Sue. "I know Big Ed. Don't let him change anything before you play and don't play him again."

"Come on, I'm no sucker. What could Big Ed do?"

The week passed slowly for The Victim, but he thought it would give too much away to request an earlier date. "I'm ready," he announced to Big Ed.

"So am I," said Ed. "Do you want to go first or second?"

"First," said The Victim gleefully.

"And I'll double the bet if you protect me from my unlucky number, 11."

"What does 11 have to do with anything?"

"My granddad was shot in Abilene after rolling an 11 in a game of craps and it's been a family curse. What do you say? A $500 bet if you protect me from 11."

"OK, I won't make 11."

"OK. You won't make 11 and at no time will the sum of the cards you alone have turned be 11."

"I agree."

Ed scribbled on a piece of paper: "I agree not to play so that I make the number 11 at any time in the game either with my own cards alone or as a sum of all cards played. Ed agrees to pay me $500 if I win; I will pay him $500 if I lose our next game of 31 (rules attached)."

"Read the rules and initial them and then sign the agreement. See, I already signed it at the bottom 'Countersigned by Big Ed.' "

The Victim signed. "Now can we play?"

"Sure thing," said Ed, grinning from lobe to lobe.

The play proceeded: Victim 5, Ed 5 (making 10), Victim 2 (making 12), Ed 5 (making 17), Victim 2 (making 19), Ed 5 (making 24). The Victim reached for the 2 to make his planned total of 26 with no 5s left. "No can do," said Ed.

"Why not?"

"Eleven protection," said Big Ed. "If you take a 2 it makes the sum of all the cards you took so far be 11. No can do."

Another fish had been reeled in.

③

ALG

ARITH

9.4. Arithmetic Sequences

In section 3.6 we discussed the beautiful mathematics that can result from giving a recursive definition of a sequence and seeking an explicit definition or from giving an explicit definition and seeking a recursive one.

Mr. Sperling wrote on the blackboard: "An *arithmetic sequence* is a sequence of real numbers such that there exists a real number d such that for all integers $n > 1$, $a_n = a_{n-1} + d$."

"Can $d = 0$?" asked Ayesha.

"Yes," replied Mr. Sperling. "Not very interesting, though."

Mr. Sperling continued to write. "Find a_{93} if $a_1 = -3$ and $d = 5/7$.

"That will take forever," complained Alain.

"Unless we can find a formula," suggested Tanya.

The class was divided into eight groups of three to explore the question of whether a formula could be found. None of the groups had very much success. "I wish Gauss were here," said Tanya.

"Who's Gauss?" the others asked in near unison.

"Carl Friedrich Gauss. The greatest mathematician who ever lived and no slouch as a physicist and astronomer. There was this story that he was once asked to add the natural numbers from 1 to 100 inclusive and that he almost instantly did it in his head."

"How?"

"They say he paired 1 with 99, 2 with 98, 3 with 97, and so on until he got to the 49,51 pair, giving him 4,900, and then he added 100 and 50 to get 5,050. Supposedly he was just a kid."

"So what would he have said here? 93 times we're adding 5/7 to –3."

"No we're not. We're adding it 92 times. Go to the simplest case. For a_2 we would be adding it once, so for a_{93} we would have to add it 92 times."

"You're saying the answer is $-3 + 92 \times 5/7$? What is that?"

"Use a calculator! The important thing is that we can generalize the calculation of a formula: $a_n = a_1 + (n - 1)\, d$."

"I don't believe this. You're not Gauss. How did you figure this out so quickly? Did you read ahead in the book?"

Alain looked at Ayesha and Tanya quizzically. "Actually, I saw the movie."

"Well, class," said Mr. Sperling, "has any group gotten anything to tell us about?"

"We need more time!" they all called back at once.

9.5. *Thinking Back*

Children can begin to learn about sequencing long before they learn to read. When a child learns that socks are put on before shoes, he or she has learned a sequence with two terms. Puzzle picture sequences that must be arranged in chronological order appear in picture books. We learn about sequences in a recursive format. In practical terms, we learn what to do next; in a long process we are often able to say what has come prior to or subsequent to something else. Talking about sequences explicitly requires more formal instruction.

Sequences lend themselves to accessible but important investigations. It is easy to discuss the limit of an infinite sequence intuitively and to pose questions such as whether or not the sequence formed by applying simple arithmetic operations to two convergent sequences is itself convergent and whether the limit of the new sequence can be found by applying the same simple arithmetic operations to the limits of the individual sequences. The concept of divergent sequences is easier to talk about and the divergence of specific sequences easier to prove than corresponding results about continuous functions. Most of all, the recognition of the recursive-to-explicit redefinition problem (and its reverse) opens intriguing new possibilities to students. "Predict the next term and explain your thinking" problems help students to see that alternative answers may arise when least expected and help to make students more appreciative of the difference between a convincing argument and a proof.

The possibilities are extensive. Students could be told the Gauss anecdote recounted in section 9.4 and then asked to generalize it to obtain a formula for the n^{th} partial sum of an arithmetic sequence. They might be introduced to difference equations as one means of solving recursive-to-explicit redefinition problems and might be asked at the university level to explore how some of the techniques for solving difference equations relate to certain techniques used to solve differential equations. Some questions about sequences have intuitively evident answers with elusive proofs; others have counterintuitive answers.

At the present time, not enough is being done early in formal education to mine the richness of the sequential ore. Many students would be helped to develop a love of mathematics by addressing this lack in imaginative ways.

10. Specification Without Loss of Generality

①
GEN

10.1. Three Days Before Yesterday It Was the Day After Sunday

"Three days before yesterday it was the day after Sunday," said Ms. Othmer. "What day is it today?"

Lucy looked at the calendar. "Monday."

"No, no, Lucy," Violet whined, "Three days before yesterday it was Monday."

"No it wasn't."

(The reader is urged to reread section 8.1 rather than taking sides prematurely.)

"So let's count ahead. With Monday as zero, count up to three and get to Thursday, so yesterday was Thursday. Today is Friday."

"No it isn't," insisted Lucy. "We came back to school this morning after a weekend off. Today is Monday."

"Let's go along with Lucy," suggested Ms. Othmer. "Lucy, what was yesterday?"

"Sunday."

"And what day was it 3 days before yesterday?"

"Thursday."

"Therefore, if I said that 3 days before yesterday it was the day after Wednesday, not only would it be Monday on the calendar, today would be Monday in the problem."

"That's 2 days before the day you would have said in the problem you gave to Lucy just now. Wouldn't that mean that in the problem you gave us,the answer had to be 2 days before the day you mentioned originally? Isn't that an interesting way to get the answer to your problem!" said Franklin with great joy. "I'll bet there are a lot of problems that can be solved by saying anything you want and working backward and then noticing the relationship between the answer you get and the thing you said and then mirroring that relationship in the actual problem to get the actual answer."

"You need psychiatric counseling, Franklin," said Lucy. "I have office hours this afternoon."

Ms. Othmer was thinking long and hard about what Franklin had said. After a very long pause she said, "That's excellent thinking, Franklin. Try to write it out in your journal as simply as you can. I don't want you to lose it."

"Why didn't you give us the problem with Wednesday instead of Sunday?" asked Lucy critically.

"Because Ms. Othmer wants to teach us to be better thinkers and problem explorers," said Franklin, returning his attention to the comic book he had hidden in his mathematics notebook.

10.2. The Sum of Seven Consecutive Integers

② ⑥ NUMB

"The sum of any two consecutive integers is odd," Binky said to Nick, "but the sum of any three consecutive integers is divisible by 3. The sum of the sets of four consecutive integers that I have tried are never divisible by 4, but all the sums of five consecutive integers I've tried are divisible by 5. Does this always work for odd numbers and never for even?"

Nick was part of the new mathematics mentoring program in his school. Binky was a precocious first grader who everyone had been happy to get out of their hair and into Nick's. Nick loved it.

"What made you say that it always works for three consecutive integers?"

"Because in any set of three consecutive integers, exactly one of them is divisible by 3, one is one more than a multiple of 3, and one is one less than a multiple of 3. They must add up to a multiple of 3."

Nick found it easier to think in terms of congruence. In any set of an odd number of consecutive integers, one would be congruent to zero modulo that odd number and the others would be nicely paired around zero, one of them congruent to 1, one to –1; one congruent to 2, another to –2; they would continue to fall in pairs adding up to zero modulo the odd number. If you added an even number of consecutive integers, the lack of a middle would ensure a non-zero sum modulo that even number. Binky's guess was right on target. "Why don't you try working out an argument for seven consecutive integers and then see why that argument would fail for four or six consecutive integers?"

"I did that last night. I took the septuple (5, 6, 7, 8, 9, 10, 11) and crossed out the 7 in red. Then I crossed out the 6 and the 8 in green, the 5 and the 9 in blue, and got stuck. They were on opposite sides of 7 and the less-than part canceled out the more-than part for each pair. Ten and 11 are both more than."

"Why not think 7 and 14?"

Binky looked puzzled for a second or two. "Oh I get it. Ten is three more than 7, and 11 is three less than 14. Or I could do it the other way, with 10 four less than 14 and 11 four more than 7. The multiples of 7 are all the same in this kind of problem. That's great!"

Binky went right on. "For 6 I took the 6-tuple (5, 6, 7, 8, 9, 10)."

Nick knew exactly why Binky had not used parallel verbal construction on his tuples.

"I crossed out 6 in red and then crossed out 5 and 7 in green. With your 'next multiple' trick I could now cross out 8 and 10 in blue. Nine would be left alone, three more than 6. There's no middle with even numbers; there always is a middle with odd numbers."

Nick couldn't wait for Binky to get a little older. What a great kid!

③
⑥
GEOM
ALG

10.3. Finding the Equation of a Parabola

"How do we know that half of a hyperbola isn't a parabola?" Veronica asked Leela. The mentoring concept had reached the high school level. Leela did not fit the stereotype for a mentor.

"Do you know what a parabola is?" Leela asked.

"It's a curve that goes like this," said Veronica tracing a concave upward curve in the air with her pinky.

"Not all curves that go like that are parabolas. Parabolas are very special curves."

Both young women laughed simultaneously. Leela continued: "Given a line and a point, a *parabola* is the set of points in their plane that are equidistant from the point and the line."

"I don't see it."

Leela removed a container of waxed paper from her four-dimensional pocket book and tore off a good-sized sheet before returning the container to the inside (which was considerably larger than the outside for reasons that Leela could not understand except that it had been given to her as a parting gift by a very strange, well-traveled man). She removed a stylus and a ruler from the bag and drew a long line segment on the paper. She then marked a point off the line segment and labeled it P. She folded the waxed paper so that the point touched the line segment and made a sharp crease. She moved the point slightly along the line segment and again made a sharp crease. She continued in the same way for a while. When she unfolded the waxed paper, the lines formed an envelope for a curve that was clearly shaped like a parabola.

"Do you see why each point of tangency to this illusory curve is equidistant from P and the line segment?"

"Yes. That's very good. I always thought a parabola was defined as the graph of a curve, such as $y = ax^2 + bx + c$ where a is non-zero."

"An equation can yield a graph which satisfies a definition; a definition of a curve is best understood independently of a particular choice of axes or of the orientation of the curve with respect to the axes. Think of defining a circle as a graph such as that of $x^2 + y^2 = r^2$: Would you feel that you could internalize the nature of a circle from that? Many mathematicians and students of mathematics much prefer general definitions such as 'the set of points in a plane at a constant distance from a fixed point in the plane' for a circle."

"I think I see what you mean."

"You do raise a very interesting question, however. Given the definition of the parabola, how do we find an equation for it and how do we prove that, in fact, $y = ax^2 + bx + c$ is the equation of some parabola?"

"How do you?"

"One thing is that you approach the general case by specifying several things. We take the given line as a line parallel to the y-axis, say $x = -m$ ($m > 0$), and the spe- cial point as $(m, 0)$ so that the parabola passes through the origin."

"Doesn't such specification spoil generality?"

"Not at all. Once we derive the form of the equation with those assumptions, we will be able to get the equation of any parabola by specifying m and applying at most one translation and one rotation. The specifications are made to make the initial derivation easier."

Leela removed a large pad from her handbag along with a small bag containing pens with a wide assortment of colors. "Let (x, y) be a general point on the parabola. Then, by definition, $\sqrt{x - m^2 + y^2} = |x + m|$. We'll call this step 1.

"It follows that we lose no points by squaring both sides, getting $(x - m)^2 + y^2 = x^2 + 2mx + m^2$, which simplifies with no gain or loss of points to $y^2 = 4mx$.

"That is the general parabola. Its special point, which is known as its *focus*, is at $(m, 0)$ and its special line, which is known as its *directrix*, has equation $x = -m$. The point where it turns around, known as its *vertex*, is at the origin."

"But what if that equation gives *more than* the parabola because of the squaring after step 1?"

"Good thinking. The only way extra points could have been added is if it were possible for any point to satisfy what would result from placing a minus sign on one side of the equation in step 1. No new points can satisfy that because both the square root function and the absolute value function are nonnegative."

Leela smoothed her furs. "We have now established that any parabola can be reoriented with respect to appropriate axes so that its equation is of the form $y^2 = 4mx$, the line of symmetry is the x-axis, the vertex is at distance m from the directrix."

Veronica was thinking about her date later that evening for part of this discussion. "I still don't see why we haven't lost generality by specifying where the vertex and the symmetry line lie."

Leela felt an urge to go for her knife but despite her prehistoric birth had become far too civilized to do damage to so insignificant a creature. Her voice maintained its careful modulation. "Think of it this way: You have the general parabola neatly enveloped on this piece of waxed paper. Are you not free to draw coordinate axes anywhere you want on that paper?"

A new voice, wholly unfamiliar to Veronica, responded, "We are nothing if we are not free."

"Who was that?" asked Veronica.

"Exactly right," said Leela, getting up lightly and leaping through the open window to the ground 20 feet below with perfect grace along a parabolic arc.

Veronica was left to ponder her original question with powerful new insights.

ARITH

GEN

10.4. Generalizing

Some students and their parents pay a lot of money to learn about specification without loss of generality. In some preparation courses for college entrance examinations, students are taught to posit specific answers to problems as a basis for analysis. For example, a student is asked to "find the single discount equivalent to giving an 8% discount after a 10% discount." The student is told to imagine that the item costs $100, though that information is not given. The 10% discount would lead to a price of $90, and the 8% discount on the $90 would reduce the price by another $7.20. The final selling price would be $82.80. They are now taught to say that the composite discount is 17.2%. This is a simple application of the strategy under discussion. Sometimes this strategy permits people to develop counterexamples to false generalizations or to gain insights that lead to proofs. Once again, schools have been remiss in developing skill in this important strategy in the majority of students. Opportunities for deeper insights into the nature of mathematics and the thinking needed for creative productivity in the discipline are lost.

Students are missing out on a lot of fun.

11. Systematically Accounting for All Possibilities

11.1. I Didn't Get a Head

It is traditional in American football for the team captains to toss a coin at the start of a game to determine which team kicks off and which team receives. Rip watched as the coin spun up and Lorenzo called tails.

Down spun the coin, lodging inextricably in a circular metal grate, but showing the United States Treasury Building.

"I didn't get a head," shouted Lorenzo, triumphantly. "We elect to receive."

"Hold it, Bozo," said Ox, the quarterback for the Downtown Questionable Trades High School Football team, the Convicts. "I've seen heads and I've seen tails," roared Ox, "but this is a building, and that ain't no head, and it ain't no tail."

Referee Wilbur patiently explained to Ox that some coins have buildings on the reverse instead of animals and that the definition of "tail" was the reverse side of the coin from the one with the head on it.

"Yeah, but how do we know there ain't an animal on the other side of that there coin?"

As usual, the referee prevailed. Of course, as the receiver began his runback pattern, he was hit with such a devastating tackle that the ball squirted out of his hands and was downed by a Convict. The tremor was so pronounced that it dislodged the coin. Sure enough, the coin had an abstract pattern on the side that had been hidden. What the coin was, and how it happened to be tossed will forever remain one of those mysteries that is never solved.

"We got no more objections," said Ox, laughing unattractively. "And anyway, we receive to start the second half."

Needless to say, Rip and his teammates scored not a point that day, and the team went downhill from there.

"The moral of this tale," said a local genius, "is 'Never call a toss unless you've seen both sides first.' "

Big Ed knew what he was talking about.

11.2. Solving Absolute Value Inequalities

ALG

"I find absolute value inequalities daunting," Binky admitted. Nick reflected that most second graders would.

"For example, I've been working on this one: $|4x - 3| + |7 - 2x| \leq 3$. My general strategy is to break the analysis into cases, allowing for the four key possibilities:

IA. $4x - 3 \geq 0$ and $7 - 2x \geq 0$
IB. $4x - 3 \geq 0$ and $7 - 2x \leq 0$
IIA. $4x - 3 \leq 0$ and $7 - 2x \geq 0$
IIB. $4x - 3 \leq 0$ and $7 - 2x \leq 0$

"These conditions are equivalent to

IA' $x \geq 3/4$ and $7/2 \geq x$ or $7/2 \geq x \geq 3/4$
IB' $x \geq 3/4$ and $7/2 \leq x$ or no solution
IIA' $x \leq 3/4$ and $7/2 \geq x$ or $x \leq 3/4$
IIB' $x \leq 3/4$ and $7/2 \leq x$ or no solution

"I originally found this very confusing, because the two conditions with non-null solution sets seemed to combine into the condition $7/2 \geq x$, but that is not so. They must be kept separate, because each of them leads to a different inequality when the absolute value signs are removed.

"(IA) leads to $4x - 3 + 7 - 2x \leq 3$, or $2x + 4 \leq 3$, or $2x \leq -1$, or $x \leq -1/2$. Since the set of points $\{x \mid x \leq -1/2\}$ has null intersection with the set of points determined by (IA), namely, those points x with $7/2 \geq x \geq 3/4$, (IA) leads to no solutions.

"The only remaining possibility is that (IIA) leads to solutions. (IIA) results in the given inequality becoming $3 - 4x + 7 - 2x \le 3$, or $10 - 6x \le 3$, or $7 \le 6x$, or $7/6 \le x$. Since the set of x with $7/6 \le x$ has null intersection with the set of x with $x \le 3/4$, I couldn't find any solutions whatsoever to the inequality."

"You said that you found absolute value inequalities daunting," marvelled Nick, "but where does the daunt come in?"

"I expected to have an intuitive grasp of why the inequality has no solutions, but, except for the analysis I shared with you, I see no reason to have guessed that there is not a single real number in its solution set. I so rarely have found anything counterintuitive that this daunts me."

Nick was pleased to have something to say to Binky. "What you say is not precisely so . The result is not counterintuitive, because your intuition really gave you no clue as to whether this inequality had any solutions. Perhaps a better term, although I admit it is a term of my own invention, would be that this result is *paraintuitive*, much as your ability to move objects with your brain is paranormal."

"I can't move objects with my brain, Nick."

"Yet."

"Now just this morning, I started working on $|8x^2 + 2x - 3| + |7 - 2x| \le 3$. Can we talk about that one?"

"Next time," said Nick, with a curious mixture of inexplicable emotions.

④ ⑬ ALG

11.3. Finding the Counterfeit Coin

"I have 12 coins, 11 of which are identical and true, and one of which is an insidiously clever counterfeit," said Professor Moriarty to his freshman mathematics class at Cambridge University. "It is your job to find the counterfeit in the fewest possible weighings."

"Oh, no," said young Van Dusen. "The old balance scale problem. How boring."

"You arrogant pup!" said Moriarty, accurately reading Van Dusen's mind. "You're no Thinking Machine yet. Actually, you don't have a balance scale. One of my colleagues, a Frenchman named Dupin, has invented the most accurate scale, which yields weights correct to a ten-millionth of a gram, and I assure you that the difference in weight between a true coin and the counterfeit exceeds the tolerance of the scale. Since you find this so simple, would you please explain to the class how many weighings you require and exactly how you will proceed?"

Two students in the class actually snickered. One was a Chinese young man who had enrolled under the name of Ming, and the other was a promising young Swiss science major named Sivana. "An interesting conundrum," said the Chinese gentleman.

"Van Dusen would need the wisdom of Solomon to solve this one," thought Sivana.

"A truly curious problem," said the one known as Ming. "If we had but one coin and were ensured a counterfeit with the set, we could identify the counterfeit with no weighings; yet if we had two coins, the problem would be unsolvable."

Sivana picked up the chain of reasoning with some delight. "The simplest productive case to contemplate is the one with three coins. It is obvious even to Van Dusen that three weighings would suffice. But perhaps it is an open question to him whether it can be done in fewer."

"It is eminently clear that it cannot," said the one called Ming. "One cannot gain knowledge of the weights of all three in two weighings, and knowing the weights of two would leave the question of which is counterfeit unsolved, were the two of different weights."

"How far could one go with three weighings?" asked Sivana, rhetorically. "I conjecture that we may be able to do as many as five."

"Seven," said the one called Ming. "Recognize that in two weighings, knowing the weight of a true coin, one could determine the one that is counterfeit among four coins with certainty. With one weighing, knowing the weight of a true coin, one could do the same thing with at most two coins. In general, in n weighings, after we know the weight of a true coin, we can find the counterfeit among a set 2^n coins by dividing successively by 2, always limiting our attention to the set known to contain the counterfeit.

"With seven coins, we might productively begin by weighing four coins on the first trial, and then two of them with one of the three remaining coins on the second trial. Let us say that the weights are, in order, p and q. If $q/3 = p/4$ then all five coins which have been weighed are true, and we can compute the weight of a true coin, so that on the third weighing, by weighing the sixth coin alone, we can determine whether it is the sixth coin or the seventh coin that is counterfeit."

"But what if $q/3 \neq p/4$? How could we possible determine the counterfeit from among the five candidates in one weighing?"

"Let us call the first four coins A, B, C, D, the fifth coin E, and the other two coins F and G. The combined weight of A, B, C, and D we have called p, and the combined weight of A, B, and E is q. We have assumed that $q/3 \neq p/4$, so that we know that one of the five coins A, B, C, D, E is the counterfeit. On the third and final weighing, we can weigh A and C together obtaining the weight r. We proceed by breaking into cases, systematically accounting for all possibilities.

"If $r/2 = q/3$, we deduce that all of the coins A, B, C, E are true and, therefore, D is the counterfeit. If $r/2 = p/4$, we deduce that E is the counterfeit. If all three weighings are incompatible (i.e., if $r/2 \neq q/3$, $r/2 \neq p/4$, and $q/3 \neq p/4$), we must resort to algebra.

"Let w be the weight function. We solve simultaneously the equations

$$\begin{aligned} &w(A) + w(B) + w(C) + w(D) = p \\ &w(A) + w(B) + w(E) = q \\ &w(A) + w(C) = r. \end{aligned}$$

"Under each of four cases, using t to represent the weight of a true coin and x the weight of a counterfeit:

Case I: A is counterfeit
$x + 3t = p$
$x + 2t = q$
$x + t = r$

This would lead to $t = q - r$ and $x = 2r - q$. We see whether both of these quantities are positive and whether $3t + x$ does equal p. If so, Case I pertains, and A is the counterfeit.

Case II: B is counterfeit
$x + 3t = p$
$x + 2t = q$
$2t = r$

This would lead to $t = r/2$ and $x = q - r$. Again, we see whether both of these quantities are positive and whether $3t + x$ does equal p. If so, B is indeed counterfeit."

Sivana interrupted. "What if cases I and II both lead to consistent results?"

The one called Ming smiled. "But, dear Sivana, consider the value of p in each case. In case I, $p = 2r - q + 3(q - r)$, or $2q - r$. In case II, $p = q - r + 3(r/2)$ or $q + r/2$. Thus $2q - r = q + r/2$, or $q/3 = r/2$. But this result violates our assumption that the first two weighings were not consistent!"

"I'll be a mad scientist!" cried Sivana, "And I'll wager as we carry this through, the same type of thing will continue to occur."

"Let us test your wise wager.

Case III: C is counterfeit
$x + 3t = p$
$3t = q$
$x + t = r$
$t = q/3 \qquad x = r - q/3$
$p = r - q/3 + 3(q/3) = r + 2q/3$

"Let us see if this case can be consistent simultaneously with a consistent case I.

"From case I, $p = 2q - r$; from case III, $p = r + 2q/3$; but if $2q - r = r + 2q/3$, then $2r = 4q/3$, or $r/2 = q/3$, violating the assumption of inconsistency of the second and third weighings.

"Similarly, from simultaneously consistent cases II and III, we would be driven to $q + r/2 = r + 2q/3$, or $q/3 = r/2$, the same violation.

"Let us consider

Case IV: E is counterfeit

$$\begin{aligned} &4t = P \\ &2t + x = q \\ &2t = r \\ &t = r/2 \qquad x = q - r \\ &p = 2r \end{aligned}$$

"We must now rule out the possibility of simultaneous consistency of this case with each of the previous three cases:

Cases IV and I

$p = 2r = 2q - r$, or $3r = 2q$, a violation of the inconsistency of the second and third weighings

Cases IV and II

$p = 2r = q + r/2$, or $3r/2 = q$, or $3r = 2q$ again.

Case IV and Case III

$p = 2r = r + 2q/3$, or $r = 2q/3$, or $3r = 2q$ yet once again."

"One thing remains," said Sivana. "How do we know that it would not be the case that none of the four leads a consistent solution?"

"That violates either the given information that there is exactly one counterfeit coin or our previous deduction that the counterfeit had to be either A, B, C, or E, because of the inconsistency of each pair of three weighings."

Professor Moriarty had approached them. "Have you decided how many weighings are need for 12 coins?"

The one called Ming looked up with the expression that a teacher would have in contemplating a brilliant student rather than what seemed to be the actual situation. "Dear Professor Moriarty, that number 12 was a snare. You could do as many as 15 in four weighings. In fact, in n weighings you could find the counterfeit from among as many as $2^n - 1$ coins. I'll mail you a general proof of that next month."

"Why will you have to mail it?" Moriarity asked.

"A slip of the tongue," said the one called Ming, dismissing Moriarity.

Sivana contemplated his classmate closely. "That was no slip of the tongue. Would you share the truth with me?"

"I am simultaneously enrolled in three doctoral programs. As you know, I am enrolled in a Doctor of Philosophy program at this university. I am also a candidate for a Doctor of Medicine at the Sorbonne and for a Doctor of Laws at Harvard University. Soon you may call me Doctor . . . "

"Doctor Ming?"

"No. There will be another name."

11.4. The Cases Closed

The analysis of problems through cases can be exceedingly complex, as the reader will find who attempts the proof in the case of n coins one of which is counterfeit, as the insidious Doctor promised Moriarty in section 11.4. When a coin is tossed, the analysis by cases is somewhat simpler, even when the coin toss results in violence because the coin is two-headed. Young children can be introduced to proof by cases easily and enjoyably in practical situations. Surely, by the time they have finished high school, every student should be comfortable with this productive problem-exploration strategy.

The additional hypotheses that are available in each section of an analysis or proof by cases make it easier to complete the arguments. Sometimes we need all the help we can get.

12. Using a Computer

12.1. A Polynomial Formula for Primes

(2)
(7)
ALG
NUMB

"There is no such thing as a polynomial formula for primes," said Tasha, looking up at her younger brother from her personal computer.

"I have one," Franklin insisted. "Consider $p = n^2 + n + 41$. Try it. You'll see."

Tasha sneered as she began substituting negative numbers first to discomfort her sibling.

"When $n = -3$, $p = 9 - 3 + 41 = 47$. Her sneer lessened as she substituted -2, -1, 0, 1, 2, and 3, obtaining a prime in each case. "This can't be."

"It is," said Franklin, grinning broadly. "I'm going to be famous." He left the room without closing the door, to Tasha's growing annoyance.

Tasha knew her strength lay in programming and systems analysis much more than in mathematics, but she knew that any bold numerical conjecture must be tested extensively before attempting a proof. She put in the disk on which her program for testing integers to find out whether they were prime or composite was stored. She wrote a program to test all integers of the form $n^2 + n + 41$ for $n = 1$ to $n = 100{,}000$ for being prime or composite, using her disk program as a subroutine and instructing the computer to print out the least n for which $n^2 + n + 41$ is composite. In a few minutes, the program was ready to run. One second later, she heard the printer spewing something out. When she contemplated the printout, her face reddened.

"Hey, twerp!" she screamed through the open door. Franklin bounded in. "That's no way to speak to your genius brother."

"Well, genius brother, how can $n^2 + n + 41$ be prime if the same integer greater than 1 is a divisor of all three terms?"

"Huh?"

"Did you ever think what would happen if you let n equal 41?"

Tasha's best friend Lara was thinking of going into teaching. She had often expressed doubts about the value of computers in the mathematics classroom for reasons Tasha could not understand. "Look, Lara. Working by hand, it would be most tiresome to check numbers resulting from substitutions from $n = -40$ to $n = 40$, each of which turns out to be a prime. Unless a student saw the fact that 41 would yield a composite, the student's case-by-case attempt to test the conjecture that $n^2 + n + 41$ always yields a prime would be tedious. The computer is a powerful adjunct for testing conjectures over a range of numbers far beyond the capacity of humans."

Lara was sharp, flexible, and sensitive. "You know, Tasha, you're right."

12.2. Finding Primitive Pythagorean Triples

GEOM

ALG

"Other than 3-4-5, 5-12-13, and 7-24-25, are there any other primitive Pythagorean triples?" Mary asked Binky.

Although Binky was only 9, he had developed a crush on the little red-headed second grader who her teacher was touting as a roaring genius, based on 3 months of teaching experience.

"Do you know how to program a computer yet?" Binky asked.

"Sure," Mary said. "My mom taught me last year."

"Write a program generating triples (p, q, r) for all ordered pairs of integers (m, n), $0 < m < n$, by the algorithm $p = n^2 - m^2$, $q = 2mn$, and $r = m^2 + n^2$, and print all such triples for which p and q are relatively prime."

"What does 'relatively prime' mean?"

Binky felt a warm glow through his whole body. "Two integers are relatively prime if they have no common prime factor."

Mary thought for a moment. "So 8 and 9 are relatively prime, even though neither of them is prime."

"You've got it," said Binky.

The next day, Mary, with a little bit of help from her mom, came to school with four pages of densely packed printout of primitive Pythagorean triples. In addition to standards such as 8-15-17 and 9-40-41, her list included such unfamiliar triples as 177-1736-1745.

"My mom says that all triples of the form you gave are Pythagorean, though not necessarily primitive. I really like this list. I'm gonna have a lot of fun with this in school."

"Not for a few years," said Binky. "They don't teach algebra in second grade. You'll have to wait until fourth grade in this school. Then you'll have fun."

"I wonder if I would have had as much fun with this list if I knew the proof."

"Sure you would. The computer helps me to see that even the most abstract proofs have real meaning. It helps to make mathematics real."

GEN

12.3. *Learning Chess*

Mary wanted to learn chess as quickly as possible so that she could have the pleasure of playing Binky on fairly equal terms. She asked her father for a chess program with many levels of skill. The program had 15 levels, ranging from "We Just Move" to "If You Beat Us Now, You Ought to Be Ranked." Within 1 week, Mary had gotten to Level 6, which was called "You Can Probably Beat Your Father." She decided that the program had been written by a woman, because Level 7 was called "You Can Probably Beat Your Mother."

Binky was genuinely surprised that it took him 16 moves to mate Mary. Most opponents he played fell in under 12 moves, and Mary had just learned the game a couple of weeks ago.

"I love learning chess on a computer," said Mary. "I turn it on when I feel like it, I can take back moves, it never laughs or gloats, I can turn it off anytime I want and decide whether or not to save the games so far. I can choose the level on which to play, and I can take the things I learned and write about them during or after the play. I think I could learn anything if I didn't have to deal with people saying, 'Very good, Mary, but I suggest you talk it over with some of the other kids in your group before entering it in your journal.' "

"Someday," Binky predicted, "computers using games, simulations, and multilevel instructional density will be commonplace."

"I don't understand what you mean yet, but when I do, I'm sure I'll agree with you."

Binky reflected that Mary wasn't a roaring genius in math but, in his own words, she was plenty good.

⑦

12.4. *Refining Conjectures*

NUMB

Mary had reached fourth grade without getting to use her sheaf of primitive Pythagorean triples, but a wonderful new opportunity presented itself when Mrs. Leeds said, "Between any two fractions, there are infinitely many fractions."

"She meant rational numbers, Mary. Teachers often confuse numbers and numerals, but nowhere is this more confusing to students than when they're studying rational numbers. A fraction names a rational number. It is a symbol, specifically a numeral. I mean, how many letters does Mary have?"

"Four," Mary replied, in a perplexed tone.

"Let me see them," said Binky.

"Can I borrow a piece of paper?"

"Why?"

"To write my name."

"I didn't ask how many letters there were in your name. I asked how many letters Mary had."

Mary was still perplexed. Binky took out a piece of paper and wrote a minuscule numeral 5 in the lower left-hand corner and a monstrous numeral 2 spanning the rest of the sheet. "Which represents the larger number?"

It all hit home at once for Mary. Of course 5 was a bigger number than 2, and of course she had mailed the letter her mother had given her this morning before she had got to school.

"But if Mary has no letters, how would you ask me about the number of letters in my name?"

"I'd use your name's name. Mary is a girl (oh yes), but 'Mary' is a girl's name, and has four letters. Mary had one letter when she left her house, Mary has no letters now, but 'Mary' always has four letters."

Mary went up to the blackboard and wrote, "One half is a rational number, as is 1/2, but '1/2' is a fraction. Also, '0.5' is neither a rational number nor a fraction, but a decimal, yet 0.5 = 1/2 is a rational number. Finally, '0.5 = 1/2' is an equation."

"Brava!" said Binky.

"So," continued Mary, "if two rational numbers are named by fractions, there are infinitely many nonequivalent fractions naming rational numbers between them."

"You ought to be on videotapes," said Binky.

"So I was wondering, Binky, how can you predict the fraction with least denominator that names a rational number between two given rational numbers?"

"Let's go to the computer," said Binky.

After about 20 minutes, they were running a very nice program that permitted them to enter an ordered pair of ordered pairs, each component of which consisted of relatively prime natural numbers, and to obtain the ordered pair which, when used to name a fraction by writing the simple numeral for its first component before a slash line followed by the simple numeral for its second component, was the fraction with least denominator naming a rational number between the rational numbers named by the fractions generated by the two input ordered pairs, when the process described above was applied to each of them.

(Note: The previous accurate but nearly unreadable sentence is not recommended for classroom instruction. It is far better to illustrate the meaning with a couple of samples, such as:

Input	*Output*
((1,3), (1,2))	(2,5)
((2,5), (1,2))	(3,7)
((1,4), (8,9))	(1,2)

It is usually better to present complex concepts through example rather than through verbalization, as the reader has unfortunately discovered.)

"This kills my conjecture," said Mary.

"This kills my conjecture also," said Binky.

"I'll show you my conjecture if you show me yours."

"OK," said Binky.

"I thought that you got it by taking the average of the two rational numbers and reducing it, but that doesn't even work with 1/2 and 1/3, since 2/5 beats 5/12."

"I thought you got it by adding numerators and denominators, which works when the fractions are close in some sense but fails with 1/4 and 8/9. I guess our little search-and-try program tells us that we both have to go back to the drawing board. Computers are very good for that. They help us to confirm valid conjectures and refute poor ones."

"Have you got any guess at all as to when the procedure of adding numerators and denominators works?"

"Actually, I do," Binky said. "It always works when the cross products of the two given fractions differ by one. In fact, the new fraction retains that same property with each of the original pair, so you can make a chain."

"That's exciting," said Mary.

Binky smiled. He always smiled when he had something to prove.

12.5. REM Statements

Computers should be tools in every classroom, not only the mathematics or science classroom. Access to information, word processing, data management, testing and refuting conjectures, searching for patterns, and other effective problem-exploration techniques are greatly helped by the use of our electronic friends. Computer literacy is a toddling baby step; to unleash the power of this relatively new technology, computer intimacy must be fostered in all children as early as you can get their little fingers on a keyboard with purpose.

13. Deductive Reasoning

⑧

13.1. Syllogisms

LOGIC

GEN

"If you let me have a bite of your peanut butter and jelly sandwich, I'll let you have a bite of my croissant," said Mary to her friend Michelle.

"If you let me have a bite of your croissant, I'll let you have a bite of my peanut butter and jelly sandwich." Michelle paused for a minute. "So give me a bite of your croissant."

"Why?" said Mary. "You haven't given me a bite of your peanut butter and jelly sandwich yet. Somebody has to do something before any promises need be kept."

"I'm not going to give you a bite of my peanut butter and jelly sandwich," said Michelle.

"I'll let you have a bite of my croissant anyway," said Mary.

Michelle took a bite. "Now you're a liar," she said.

"No, I'm not," said Mary, "I haven't broken any promises. By the way, are you going to let me have a bite of your peanut butter and jelly sandwich now?"

"No," said Michelle, "I told you I wasn't going to."

"Then you're a liar," said Mary. "You broke a promise."

"How could I be a liar if you're not?" asked Michelle.

"Because I didn't do the 'if' part of anything that required me to do a 'then' part. But I did the 'if' part of what you said, so you're a liar for not doing the 'then.' 'If-then' only obligates us to do something when the 'if' happens."

"I never understand you, Mary," said Michelle. "You always make me feel angry."

"So tell me, Binky, what is this all about? If I know 'if p then q' and I also know 'if q then r,' then the Law of Syllogism lets me know 'if p then r,' but that doesn't really tell me anything unless I somehow know 'p.' You could also know 'not r,' which would let you deduce 'not q' and 'not p.' But mathematics always seems to be about deducing things from if-then sentences. How will we ever know anything?"

"Mathematics can be done in a black box, without reference to reality, but in order to know something about reality, we must have a peephole in that box. Without real-world input, we can have the most elaborate system, and yet have no real-world output."

"So a syllogism without anything independently known or given is just a sillygism."

"You have a way with words, Mary."

13.2. Deduction by Mathematical Induction

④

LOGIC

"Doctor Süss was talking about deductive and inductive reasoning today in our debate seminar. Some of the students who are also taking fourth-year mathematics said that they knew about inductive reasoning from their study of mathematical induction. I said that mathematical induction is deductive, and they all started laughing. I'm right, am I not, Dad?"

Jerry beamed at his son. "Of course you are. I think it a misnomer to call inductive reasoning 'reasoning.' You are not really reasoning when you do inductive reasoning. You're making educated guesses."

"Which doesn't exactly tie into something that I've been meaning to ask you about. We have a new guy in the class who transferred from some foreign school.

His name is Allard. He gave me a problem involving mathematical induction that I find puzzling."

"The billiard balls?" asked Jerry.

"No, that's an easy one. I figured that out once I began to formulate my inductions precisely. I think of the law of natural induction as follows:

> If $p(n)$ is an open sentence about natural numbers such that
>
> 1. $p(1)$
> 2. for all $m \in \mathrm{N}$,
>   if $p(m)$ then $p(m+1)$
>
> Then (*) for all $n \in \mathrm{N}$, $p(n)$.

"In that old billiard ball problem, we are attempting to prove that any set of solid-colored billiard balls must consist of balls all of the same color. The fallacious attempt at natural induction, which I prefer to say to the common 'mathematical induction,' goes as follows:

- Anytime I have a set of one billiard ball that is solid color, every ball in the set has the same color (establishing condition 1)
- By the inductive hypothesis, we may assume that any set of m solid-colored billiard balls are all of the same color. You hand me a set of $m + 1$ billiard balls that you assure me are each solid colored. I keep my eyes closed and put one of them in my large left pocket. Since I now count m balls on the table, I know they are all of the same color, which, for purposes of clarity we may call 'carmine.' I now, still with my eyes closed, take one of the carmine balls from the table and place it in my large right pocket. I remove the solid-colored ball of unknown color from my left pocket and place it on the table with the other balls. I again have a set of m solid-colored balls, so I know they are all of the same color. Since those remaining on the table were carmine, they must all be carmine. Thus, when I remove the known carmine ball from my right pocket and place it on the table, I am guaranteed that all $m + 1$ balls are carmine (completing the proof of condition 2)
- By natural induction, we may now deduce that, for all natural numbers n, any set of n solid-colored billiard balls are all of the same color.

"Which is a bit troubling, since just yesterday I saw a set of three billiard balls, two of which were solid black and one of which was solid white. Of course, as I said earlier, this 'proof' is fallacious, because the proof of condition 2 fails when $m + 1$ equals two. Clearly, if there were a set of two balls on the table, and I put one in my left pocket and the other in my right pocket, there would be none left on the table to ensure the commonness of color. This little problem taught me how careful we have to be in the proof of condition 2, to see that there is no value of m for which our argument breaks down."

"Your discussion was astute, but I think it would have been clearer had you specified that you were taking $p(j)$ as the sentence 'Any set of j solid-colored billiard balls all must have the same color.' When you're being so careful, you should not assume that your listener is willing to supply the information you have omitted."

"I always expect something good from my listener, especially when it's you or Mom."

"So tell me, Nick, what did Allard puzzle you with?"

"He said: 'Certain descriptors define natural numbers. For example, *the least natural number* defines *one*. Suppose we specify some existing dictionary as the standard dictionary. I ask you to consider the set S of all natural numbers which can be defined by descriptors of 25 words or less, each of which appears in the standard dictionary.

" 'Any dictionary, and specifically our standard dictionary, must contain a finite number of words. There will therefore be only a finite possible number of descriptors of 25 words or less, each of which appears in the standard dictionary. Thus S is a finite set.

" 'We know that the natural numbers are *well-ordered*; that is, we know that any nonempty set of natural numbers has a least element. In fact, the well-ordering principle is formally equivalent to the principle of natural induction. Because of the finiteness of S, its complement S' is non-null. Therefore, by well-ordering, S' has a least element, which we may call z.

" 'Thus z is *the least element of the natural numbers that has no descriptor of 25 words or less, each of which appears in the standard dictionary.*

" 'Assuming that we have taken a standard dictionary of reasonable size, it will contain each of the words in the descriptor above, yielding a gruesome paradox. We have found a descriptor of 25 words or less, each of which appears in the standard dictionary, for a natural number that, by definition, can have no such descriptor.'

"This guy Allard is weird. After he presented this paradoxical situation, he stood there laughing a weird hollow laugh that gave every one of us the chills."

"Do you see how to resolve Allard's paradox?"

"No. Do you know anyone who knows?"

Polly Sue's voice drifted in from the kitchen, where she was preparing a delightful snack of fresh morels sautéed in butter. "Give a careful definition of both 'descriptor' and 'define,' and there will be no paradox."

"Do you understand what Mom means?" asked Nick.

"Let's see. If you allow descriptors to be random sequences of words and mean by 'define' to establish a one-to-one correspondence between these random sequences and natural numbers, there will be no paradox, because the descriptor that was put forward as creating the paradox will have already been used. If, on the other hand, you insist that descriptors have meaning, then definitions must uniquely specify natural numbers within the language of our mathematical system. In this case, the descriptor that was put forward as creating the paradox does not define a natural number, because it uses the word 'descriptor,' which is neither a prespecified undefined term nor a defined term in our mathematical system. Since it is not a definition, there is no paradox.

"Allard's discussion gets at critical issues in mathematical logic. You may want to bone up on that discipline for your own edification. I'd like to meet Allard sometime."

"If you guys don't hurry," warned Polly Sue, "there won't be any left."

13.3. *The Hat Trick*

Old King Cole's grandfather had been a slick cookie. He admired mathematical acumen and had decided that his beautiful daughter should marry the best logician in the country. He had invited the three best logicians he could identify in his kingdom to a timed contest that would determine who would be given his daughter's hand in marriage. He had told the three men that he had a box containing exactly three red hats and two white hats and that he would blindfold them and place a hat on each man's head; he would then direct his chamberlains to remove the three blindfolds simultaneously, so that each man could see the color of the hat on each other man's head but not the color of the hat on his own. The first man who could correctly identify the color of the hat on his own head through logic would win the hand of the princess if he could explain that logic flawlessly.

An incorrect guess or faulty logic would be penalized by 30 years of servitude in the stables.

Old King Cole's grandfather had a commitment to absolute fairness, so he decided to symmetrize the problem by placing a red hat on the head of each of the three logicians. Thus it was that when the chamberlains simultaneously removed the three blindfolds, each of the three men saw two red hats.

Cole's grandfather said, "I am ready to hear a cogent explanation but beware—the shovels are ready."

Jason approached the throne and spoke as follows: "I said to myself, 'Suppose I were wearing a white hat. Each of the other two men would be seeing one red hat and one white hat and would know that their own hat could not also be white, because the other contestant would be seeing two white hats and would instantaneously know that his own hat must be red.' Since neither of my esteemed opponents spoke within a few seconds, I realized that my hat must be red."

Old King Cole knew that this problem was old hat to even the lowliest peasant in his domain, let alone to the logicians and scholars. Thus it was that his own problem seemed to have no solution. He had a handsome, young, brainless son, for whom he wished to find the most intelligent maiden in the kingdom by a process as much like his grandfather's as possible, yet in some important way new. His scouts, knights, and fiscal agents reported to him that in fact his kingdom included 17 incredibly beautiful and amazingly intelligent young women who were stars in the world of logic. He invited all 17 to his court, but he was not at all sure what he was going to do with them.

"Can't I keep them all?" his son asked. Old King Cole realized that he had better lock his son in the east wing of the palace until the selection process was completed.

Cole consulted with Professor Silbert, who had been his father's teacher and then his own. "What should I do?" he asked.

"Do as your father would have suggested, if he had lived," Silbert replied. "Tell the young ladies that you have a large quantity of red hats, but only 16 white ones and that you will do much the same thing that your grandfather did, by putting a hat on each of their heads while they are blindfolded and then having the blindfolds simultaneously removed, with the prize that the first young lady correctly to identify the hat color she wears will be the bride of the prince if she can cogently justify her reasoning. The penalty for error or noncogency will be 17 years servitude in the scullery."

Old King Cole wasn't sure exactly how Silbert's suggestion would work, but it was far better than anything he could think of, so he carried it out exactly as suggested. He watched impatiently as the young ladies were gathered, blindfolded, hatted, and simultaneously unblindfolded.

Minerva of the Double Helix was by far the wisest lass to have graced Cole's kingdom in many generations. She knew that Cole was going to put red hats on everybody before she opened her eyes to observe 16 red hats. "Telling him, 'Woman's intuition' will earn me hard time in scullery," she thought. "If only this had been a simpler case such as the traditional three, or even four. I could do it with four contestants and three white hats by saying, 'If I had a white hat, the other three smarties would see that the problem reduced to the traditional one and solve it in a flash, because they can leave me and my hat out and just look at each other.' "

"Holy moley," shouted Minerva, raising her hand just before Athena of the Triple Catenary.

"You may approach the throne," said Old King Cole merrily.

"Will a mathematician be present to judge the cogency of my solution?"

"I know everything," said a voice behind the throne. "I am Silbert."

"I have an elegant inductive proof that leads to the inescapable conclusion that I am wearing a red hat," said Minerva saucily. "It is an induction on the following proposition: if you are in any group of n maidens, all hatted from a set containing at least n red hats but exactly $n - 1$ white hats and in which each maiden is a brilliant logician, it is deducible after the unmasking that if you perceive only red hats and no other maiden raises her dainty hand within a few seconds, then you are also wearing a red hat.

"The induction begins not with $n = 1$, but with $n = 3$. Will you allow 'classical case' as a proof for $n = 3$?"

"Of course," said Silbert, already in love with Minerva of the Double Helix.

"Let us then proceed. Let $k \geq 3$ and let us assume (by the inductive assumption) that in any group of k maidens, the proposition holds. We now consider a set of $k + 1$ maidens. Each reasons that if she were wearing a white hat, the other maidens would eliminate her and her hat from consideration, reducing the problem to a set of k maidens. Each would raise their hands as fast as their nubile young reflexes would allow. Hence, since no hands shot upward, each can now deduce that she is wearing a red hat, and would attempt to raise her hand as fast as Athena of the Triple Catenary tried. My nubile young reflexes are stronger, faster, and better."

"And you shall have my son, Heaven protect you. You are one sharp cookie. Send a courier to tell my son that we have found him a bride."

"And how would you modify your proof for a group of n young men?" asked Silbert.

"*Mutatis mutandis,*" said Minerva of the Double Helix, blushing slightly. "And, dear future father-in-law," she added, "wait until you see what I make of your son within a year."

LOGIC

13.4. The Four Tumblers

It is evening in the great hall of the most brilliant and most feared man in all of China. A helpless victim knelt before him, begging for his life.

"I'll give you a sporting chance," said the ornately clad gentleman on the golden throne. "You need only solve a trifling problem in mathematical reasoning and your wretched life will be spared. Of course, if you fail, your death shall be slow and hideous. Do you agree?"

The wretched man stared up. "I have little choice, oh great noble Doctor."

"Bring in the table," commanded the regal figure, as five slaves leaped to obey his command. Within moments, a gold-encrusted ivory table was set before the groveling man. On each corner of the table there was a tumbler.

"Blindfold him!" commanded the man in the golden robes. "Now invert one or more of the tumblers, as your hearts dictate." His slaves obeyed.

He addressed the terrified man on the floor. "Your challenge is as follows: While blindfolded, you may touch any two of the tumblers at the corners of this ivory table to determine their orientation; you may then change the orientation of none of them, one of them, or both of them; if all four tumblers, after your move, are oriented in the same way, all top up or all bottom up, the great gong of Genghis Khan will sound immediately.

"If no gong sounds, you may be assured that the four tumblers are not oriented all up or all down. The table will now be spun a random number of 90-degree rotations without disturbing the orientation of the tumblers. Again, you will be permitted to touch any two tumblers, determine their orientation, and change the orientation of none, one, or both. What is the minimum number of turns that you will need to guarantee that the gong will sound?"

The men felt the hand of Death upon his heart. Clearly, the randomness of the spin made it impossible to guarantee that he would ever be able to touch the same pair of glasses on any two given turns. Perhaps that was the insidious plan. Perhaps "the minimum number" did not exist.

"I am doomed," he thought. "May I be granted 1 hour to contemplate this curious conundrum?" he said aloud.

"One hour," said the Golden Lord. "Not one second more." The great clock started to count the hour as soon as the door to the dungeon had been closed. The wretch sat on the stone floor in despair. A hollow voice sounded from nowhere and from everywhere: "Fear not, Great One, for I am here to guide you to your destiny. The trick, my son, is to mix moves where you touch adjacent tumblers with moves where you touch tumblers that are diagonally opposite. On your first move, touch two adjacent tumblers and make them face the same way (let us say, up). On your second turn, assuming the gong has not sounded, touch two diagonally opposite tumblers, and make sure they are both facing up. If the bell has not sounded, you are now assured that three tumblers are facing up and that one is inverted.

"On your third move, again touch two diagonally opposite tumblers. If one of them is inverted, you are assured that if you make it aright you will sound the gong. The only problem arises if you find them both right side up. In that case, invert one of them. You are now assured that the glasses are arranged so that two adjacent glasses are inverted and the other two adjacent glasses are right side up.

"On your fourth move, touch two adjacent glasses. There are two cases. If they are oriented the same way, you invert them both and the gong sounds. If they are not oriented the same way, you invert them both anyway, and you are now assured that each diagonal consists of two identically oriented tumblers, but that one diagonal is oppositely oriented from the other. On your fifth and final move, you touch any two diagonal tumblers and invert them both to sound the gong. The answer to the Golden Master's problem is *five*."

"Who is speaking?" asked the man whose soul had been reborn.

"I am he who fashioned the very gong that will liberate you, and you are he that will lead our land into the next century."

"Will he keep his word and let me go?" asked the radiant one, now glowing as the sun.

"He will keep his word, as you must fulfill your destiny."

The dungeon was silent. There were still 11 minutes to go, but the man who would leave the dungeon had been changed forever from the creature who had been brought in.

13.5. Proof

Proof is the connecting tissue holding the body of mathematics together, but there are many kinds of proof. Some proofs establish algorithms such as the proof which yielded the technique for sounding the gong in section 13.4. Some proofs are based on suppositions, others on powerful techniques such as natural induction. Some proofs establish the existence of a mathematical entity, while others show uniqueness. Some proofs construct an elusive mathematical object; others show that no such object may exist or that something is impossible. Some proofs work from given information to a desired conclusion, and some begin by denying that conclusion and deducing a contradiction of a result that has already been established or of a posited hypothesis. Some proofs are achieved by working backward, and some, by working both ways from the middle. Some proofs are done in one mighty stroke, and some are done by cases. Some proofs are established within months of the time that a proposition is put forward; others elude the finest mathematical minds for centuries. Some proofs are painstaking and messy, others are crystal clear and elegant.

Mathematics is a tapestry woven with many varieties of proof. Skill in all of them is among the most important mathematical exploration strategies there are. From the day that a child first realizes that the hidden face of a fair coin is a head when the face he sees is a tail to the day that he basks in the glory of a valid proof of Fermat's Last Theorem, exposure to proof and experience with proof should be a continual part of his mathematical experience.

Proof and disproof are ends and beginnings. They should be looked upon with love but not with awe.

Afterword

We have explored 13 powerful strategies for exploration of structured or ill-structured problem situations. Developing problem-exploration strategies that are explained and modeled is the route that we have chosen in this extended chapter.

Through samples and exposition, we have begun to crystallize a path to an elusive skill: that of looking into gloom and dispelling mists with light and heat.

This is only a beginning.

There exist other accessible routes for introducing problem exploration to students of all ages. Some are simple, such as deleting the final question from standard problems for solution and asking students to pose and answer as many questions as they can related to the context that has been set. Some are more open, such as teaching games with fixed sets of rules but with unspecified objectives and having students set objectives that make interesting games that they can investigate and play. It is productive in many disciplines to present real-world situations without providing specific instructions for ensuing study and guide students into appropriate applications of problem-exploration strategies to make order out of chaos. These and other methods will be explored by other authors in this volume, and perhaps by you.

We welcome you and your students to a growing community of inquiry: the world of problem explorers.

2

Unconventional Problem-Solving Strategies in Mathematics Instruction

ALFRED S. POSAMENTIER

Problem solving has been the cornerstone of mathematics instruction from the earliest times. The National Council of Teachers of Mathematics has included it as a key component of its *Agenda for Action* (1980) and its *Curriculum and Evaluation Standards for School Mathematics* (1989). It is, as a matter of fact, listed as the first standard! These latest efforts seem to have placed problem solving as a topic deserving special attention. Fitting this new role is the categorization of problem-solving techniques or strategies. What most mathematicians have for years done intuitively is now being brought to the fore in a "cookbook" type of way. This will certainly be of use to the average student, since one important tenet of problem solving is to recall similar circumstances from the past and use these experiences to solve the present problem.

For purposes of this presentation, we shall consider a few interesting strategies (See Introduction, p. viii). In each case the method was selected since it is generally a bit "off the beaten path" and not naturally expected to be part of a typical student's problem-solving approach. Students have a tendency to try straightforward methods and rely heavily on the methods shown in class. Thus more traditional problem-solving techniques frequently form the basis for their problem-solving approach. We need to show students alternate strategies so that they can build their arsenal of problem-solving tools.

Before considering some specific methods for solving problems, a consideration of the diverse thinking styles of students is in order. Experience has shown that gifted students often react to a problem situation quite differently from normal students. This is certainly not surprising. Yet we can sometimes recognize a student as "gifted" from the way he or she approaches a problem, rather than by mere achievement on classroom tests. One illustration of this is a student's reaction to the following problem:

GEOM

Find the area of the shaded region, when *ABCD* is a square of side length 1, and the two arcs are quarter circles of radius 1, with the center at opposite vertices of the square.

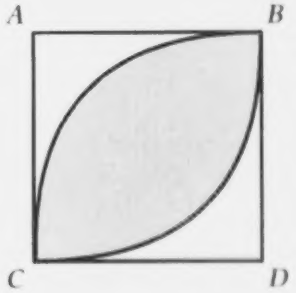


Figure 2.1

Typical mathematics coursework training often has students dissecting areas where the figure is not one of the common shapes for which a formula is known. Consequently, students may consider drawing diagonal *BD* and then finding the area of the segment shown below, by subtracting the area of the right triangle *BDC* = 1/2 from the area of the quarter circle *BDC* = $\pi/4$, and then doubling this area to get $2(\pi/4 - 1/2) = \pi/2 - 1$.

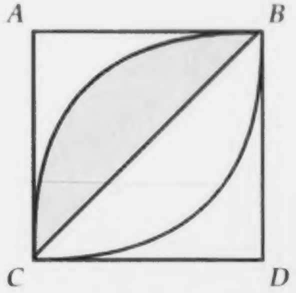
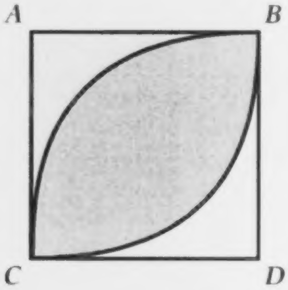


Figure 2.2

Experience shows that this is by far the most common method but not the most elegant!

A somewhat more elegant procedure, and one far more infrequently shown to be used by students, involves a "complementary approach." Here the student seeks the area of half of the nonshaded region, say the region *ABD*, by subtracting the area of the quarter circle *BDC* = $\pi/4$ from the area of the square = 1. Then doubling this area = $2\,(1 - \pi/4) = 2 - \pi/2$, and subtracting it from the area of the square to get: $1 - (2 - \pi/2) = \pi/2 - 1$.



⑥

Figure 2.3

The gifted student has shown an even more streamlined version of this method. He or she would find the area of quarter circle *ABD* and quarter circle *CDB* and deduct the area of the square from the sum of the areas of the two quarter circles. This requires "seeing" that the regions I and III are each considered once, while the region II is used twice. This taking the square from this sum leaves region II.

That is,

area of quarter circle *ABD* = $\pi/4$
area of quarter circle *DBC* = $\pi/4$
area of II = $\pi/2 - 1$.

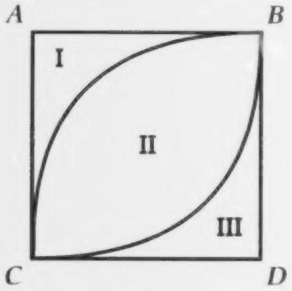


Figure 2.4

These three methods for solving a simple problem afford an opportunity for observing various thinking styles. They are shown here to sensitize the reader to various levels of thinking on the part of the student. This must be kept clearly in mind when considering the problem-solving methods discussed in the rest of this chapter. Naturally, modification of the presented material must be made by the teacher in discussing these methods with a secondary school class. Teachers play a key role in interpreting these methods in the most appropriate fashion to their students. There is no substitute for a fine teacher, especially not in the all-important topic of problem solving.

Working Backwards

The first of these problem-solving methods is called "working backwards" and is commonly used in doing proofs as well in many other situations. To show its usefulness, a few "dramatic" examples will be presented.

> If the sum of two numbers is 12 and their product is 4, find the sum of their reciprocals.

Typically, a student will solve this in the traditional way, arriving at the answer 3 after some tedious calculations. Consider the following rather straightforward and somewhat inelegant solution:

ALG

⑬

Let x = first number, and y = second number.
The key equations are $x + y = 12$ and $xy = 4$.
By substitution,

$$x(12 - x) = 4$$
$$x^2 - 12x + 4 = 0.$$

Using the quadratic formula we get

$$x = 6 \pm 4\sqrt{2}.$$

By substituting for x we get values for y:

x	$6 + 4\sqrt{2}$	$6 - 4\sqrt{2}$
y	$6 - 4\sqrt{2}$	$6 + 4\sqrt{2}$

(See also Chapter 13, p. 278)

By rationalizing denominators and simplifying,

$$\frac{1}{x} = \frac{1}{6+4\sqrt{2}} = \frac{3-2\sqrt{2}}{2}$$

$$\frac{1}{y} = \frac{1}{6-4\sqrt{2}} = \frac{3+2\sqrt{2}}{2},$$

so that

$$\frac{1}{x} + \frac{1}{y} = \frac{3-2\sqrt{2}}{2} + \frac{3+2\sqrt{2}}{2} = \frac{6}{2} = 3 \quad (\textit{Answer})$$

Using a backwards strategy, one asks, What are we looking for and how can we get back to the starting point from there?

We seek to find ①

$$\frac{1}{x} + \frac{1}{y}.$$

This could have come from the sum of these fractions:

$$\frac{x+y}{xy}.$$

Since we know $x + y = 12$ and $xy = 4$,

$$\frac{1}{x} + \frac{1}{y} = \frac{x+y}{xy} = \frac{12}{4} = 3.$$

Clearly, this backward strategy was much more elegant and efficient than the first method. A reverse strategy is certainly not new. It was considered by Pappus of Alexandria about A.D. 320. In Book VII of Pappus's *Collection*, there is a rather complete description of the methods of "analysis" and "synthesis." T. L. Heath, in his *Manual of Greek Mathematics* (Oxford University Press, 1931, pp. 452-453), provides a translation of Pappus's definitions of these terms:

> Analysis takes that which is sought as if it were admitted and passes from it through its successive consequences to something which is admitted as the result of synthesis; for in analysis, we assume that which is sought as if it were already done, and we inquire what it is from which this results, and again what is the antecedent cause of the latter, and so on, until, by so retracing our steps, we come upon something already known or belonging to the class of first principles, and such a method we call analysis as being solution backwards.
>
> But in *synthesis*, reversing the process, we take as already done that which was last arrived at in the analysis and, by arranging in their natural order as consequences what before were antecedents, and successively connecting them one with another, we arrive finally at the construction of that which was sought; and this we call synthesis.

Consider the problem alteration where $x + y = 12$ and $xy = 40$. Here the values of x and y are imaginary, which is only discovered when a straightforward method is used to first find the values of x and y. The backwards method enables a solution to the problem without first finding x and y. This may be a disadvantage in understanding the mathematics in the problem situation but certainly solves the given problem more efficiently.

A second example of this backwards or reverse problem-solving strategy shows the wide variety of application possibilities.

GEN

Using only one 5-liter and one 11-liter bucket, how can 7 liters of water be obtained in the larger bucket?

The solution requires considering the desired result and then working backwards (analysis) to the starting point. The reversal of this analysis (synthesis) will then provide a correct procedure to solve the problem.

①

- The desired result is having 7 liters of water in the 11-liter bucket.
- This leaves 4 liters empty in the bucket.
- How might that have been obtained?
- Four liters of water had to have been poured off from a full 11-liter bucket.
- To do this, we need to have 1 liter of water in the 5-liter bucket (leaving 4 liters empty).
- How can we get 1 liter of water in the 5-liter bucket?
- Pour 5 liters of water twice from a full 11-liter bucket, leaving 1 liter of water in the 11-liter bucket.
- This can than be transferred to the empty 5-liter bucket.
- Now, from a full 11-liter bucket of water, pour off 4 liters of water by simply filling the 5-liter bucket (which already had 1 liter of water in it).
- This leaves the required 7 liters of water in the 11-liter bucket.

This backwards method, when applicable, makes a rather difficult problem much simpler. It is imperative that such solution methods are shown to students at all grade levels so that they can begin to try to apply these methods when straightforward methods are either elusive or too cumbersome.

Analyzing Extremes

A second problem-solving method that may occasionally be used also requires a rather unusual way of thinking. For this problem-solving strategy, that of analyzing extreme cases, the student is required to modify variables that do not affect the problem's solution but make the problem considerably simpler. Consider some examples:

> The tangent AB of the smaller of the two concentric circles is a chord of the larger circle. Find the area of the shaded region, if $AB = 8$. GEOM

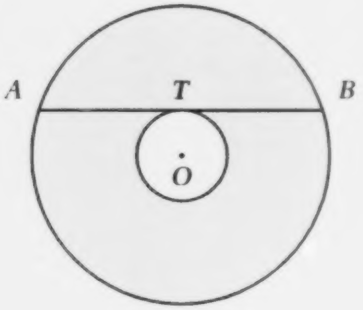
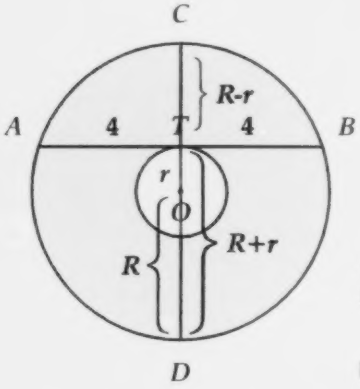


Figure 2.5

A rather straightforward solution, which, although it still uses traditional methods, becomes quite streamlined. ⑥



Area of shade $= \pi R^2 - \pi r^2 = \pi(R^2 - r^2)$
Product of Chords: $(R - r)(R + r) = 4 \times 4 = 16$
$R^2 - r^2 = 16$

Figure 2.6

Therefore, area of shade $= 16\pi$

The problem can be made much simpler by considering extreme cases of variables. The following solution, although quite novel, provides a useful procedure which enables the inspection of the problem in a changed form.

Assume the smaller circle is made very small, so small that it is reduced to a point. This can be done without loss of generality, and the problem remains intact. ⑤ The resulting situation now makes the original problem trivial and it can be solved easily. ⑩ $\overline{AB}$ then becomes the diameter of the large circle. Therefore, the area of shaded region (larger circle) is $\pi R^2 = 16\pi$.

With this technique in mind consider the following problem:

The distance between two concentric circles is 10. What is the difference between their circumferences?

GEOM

⑤

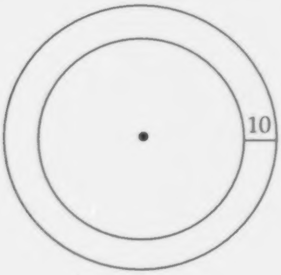


Figure 2.7

Since the size of the circles is not given, let us assume that the smaller circle is extremely small, so small that it becomes a point. Then the distance between the circles, 10, now becomes the radius of the larger (or remaining) circle. The circumference of this larger circle is then 20π, which differs from the "zero circumference" circle by 20π.

Once again this exhibits how changing variables to suit a simple solution is helpful in problem solving. The statement of the problem implied an independence from the size of the circles. This can be shown as follows.

The differences of the circumferences is

$$2\pi R - 2\pi r = 20\pi.$$
$$R - r = 10.$$

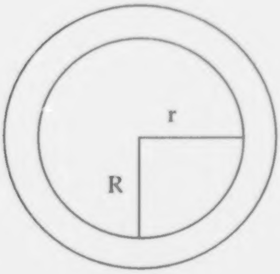


Figure 2.8

Thus the differences between the circumferences is 2π times the difference between the radii. This independence from the size of the circles leads to a very humorous problem.

If the length of a rope, stretched around the equator of the earth, is increased by 1 meter and evenly spaced around the globe, will a mouse fit under this rope?

GEOM

④

Since we just showed that the difference between circumferences is 2π times the difference between the radii, the difference between the radius of the rope circle and the earth is $\frac{1}{2\pi}$, or 0.159 meters or about 16 cm—clearly space for a fat mouse.

A third example of this technique involves a nonspecified pentagram (Figure 2.9).

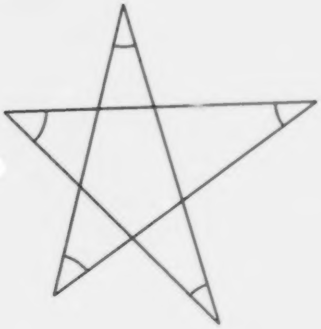


Figure 2.9a

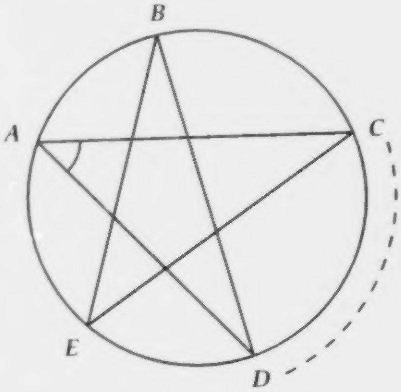


Figure 2.9b

It is known (and can easily be proved) that the sum of the vertex angles of any pentagram is the same. What is that sum? (See Figure 2.9a.)

Since the type of pentagram was not specified, we can either assume the pentagram is regular or that it is one which is inscriptable in a circle (i.e., all the vertices lie on a circle, as in Figure 2.9b). In the former case, where the pentagram is regular, all angles are congruent and a solution is readily at hand—a number of methods can be used.

In the latter case, we notice that each of the angles is now an inscribed angle of the circle and so has half the measure of the intercepted arc. Consequently, we get the following:

④

⑩

$m\angle A = \frac{1}{2}\, m\overset{\frown}{CD};\ m\angle B = \frac{1}{2}\, m\overset{\frown}{ED};\ m\angle C = \frac{1}{2}\, m\overset{\frown}{AE};\ m\angle D = \frac{1}{2}\, m\overset{\frown}{AB};\ m\angle E = \frac{1}{2}\, m\overset{\frown}{BC}$

[$\overset{\frown}{CD}$ *is* arc *CD*]

$m\angle A + m\angle B + m\angle C + m\angle D + m\angle E = \frac{1}{2}\,(\overset{\frown}{CD} + \overset{\frown}{ED} + \overset{\frown}{AE} + \overset{\frown}{AB} + \overset{\frown}{BC})$

That is, the sum of the vertices is one half the circumference of the circle, or 180°.

Again there was no loss of generality by allowing the nonspecified pentagram to assume a more useful configuration. Yet this change made the solution much more manageable (and solvable).

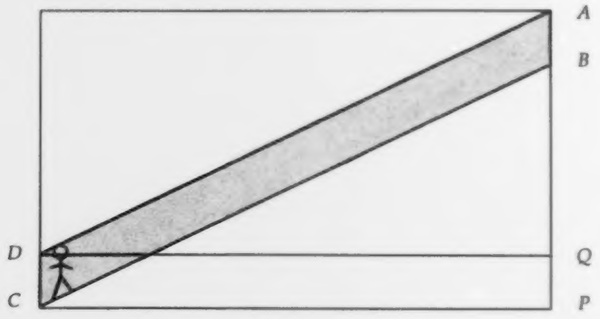
An enthusiastic student, having difficulty finding a solution to the following problem, recently asked to have his attempted solution reviewed.

GEN ⑤ Consider a rainstorm with no wind and a uniform rainfall. In order to stay driest, is it better to run through the rain from point *C* to point *P* or walk slowly, or doesn't it matter?

The student tried lots of procedures (including calculus). He was troubled by the notion that running through the rain would make his front side very wet (remember how a car windshield is affected when driving fast in a rainstorm), whereas slow walking would result in a very wet top surface. He tried to adjust these variables. He needed to determine a key question—this time, one having to do with extremes. What is the effect of the person's *time* in the rain? If one ran extremely fast (say, almost infinitely fast) through the rain, he would spend far less time in the rain than if he were to walk extremely slow (say, almost no movement) through the rain and get drenched in the process. The problem is thus reduced to one more of reasoning (using extreme cases) rather than one which is computational.

The original problem is actually split into two subproblems: (a) What effect does the speed have on the top of the head? and (b) what effect does the speed have on the front of the body? Using the *analysis of extremes*, the effect on the top of the head is clear: At very slow speeds, lots of rain will hit the top of the head, while at very fast speeds the top of the head would be exposed for less time, thus reducing the amount of rain to hit the top surface. The analysis of extremes is just a bit more involved for the consideration of the front of the body. To enable a more simplified discussion, consider the vertical rain falling uniformly at the same rate of speed as the person is running from C to *P*.

⑥



The raindrops in *AB* are the last he will hit when reaching *P*. The parallelogram *ABCD* represents (in two dimensions) the raindrops he will come into contact with throughout his journey from *C* to *P*. Were he to go infinitely fast, he would sweep out an area (again, in two dimensions) of the rectangle *QPCD*. Part b of the problem is then reduced to comparing the area of *ABCD* with the area of *QPCD*. Since they share the same base *CD* and have the same altitude *PC*, their areas are equal.

Thus, for the case where the rain is falling uniformly (at the speed of the runner) and without wind, the runner's speed does not affect the front surface, only the top. Therefore, traveling faster in the rain makes for an overall drier person!

The analysis of extremes, used with success in this problem, is a very powerful source for asking a key question. Many everyday (nonmathematical) problems sometimes can be more easily analyzed by posing a question about the extreme situations. For example, when a question in economics is posed, a consideration of the behavior of the variables under the extreme conditions (both high and low) is useful in understanding and ultimately answering the question.

Simplifying the Problem

Changing variables need not always involve extreme cases; sometimes a problem's solution can be simplified by other changes. The following example shows how making the variables "more desirable" renders a trivially simple solution to an otherwise more complicated problem.

Find the sum of the coefficients in the binomial expansion of $(x + y)^8$. (The solutions of this problem also demonstrate the power of intuition, insight, and experience.)

ALG

In the traditional, straightforward method,

$$(x+y)^8 = \binom{8}{0}x^8 + \binom{8}{1}x^7y + \binom{8}{2}x^6y^2 + \binom{8}{3}x^5y^3 + \binom{8}{4}x^4y^4 + \binom{8}{5}x^3y^5 + \binom{8}{6}x^2y^6 + \binom{8}{7}xy^7 + \binom{8}{8}y^8.$$ ⑬

The sum of the coefficients is

$$\binom{8}{0} + \binom{8}{1} + \binom{8}{2} + \binom{8}{3} + \binom{8}{4} + \binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8} = 1 + 8 + 28 + 56 + 70 + 56 + 28 + 8 + 1 = 256$$

For the second solution, let $x = y = 1$ in the first equation of the above solution. Then

$$(x+y)^8 = (1+1)^8 = \binom{8}{0} + \binom{8}{1} + \binom{8}{2} + \binom{8}{3} + \binom{8}{4} + \binom{8}{5} + \binom{8}{6} + \binom{8}{7} + \binom{8}{8},$$ ⑩

so that the sum of the coefficients is $(1 + 1)^8 = 2^8 = 256$.

The way a problem is physically presented and then rearranged (or modified) for solution is another way of changing variables (without disturbing the generality implied in the problem statement) to facilitate solution.

Consider the following:

GEN

GEOM

⑥

A bug perched on corner B of a solid box chooses to reach the diagonally opposite corner H by the shortest route along the faces of the box ($AB \neq AE$.) Which path does it take?

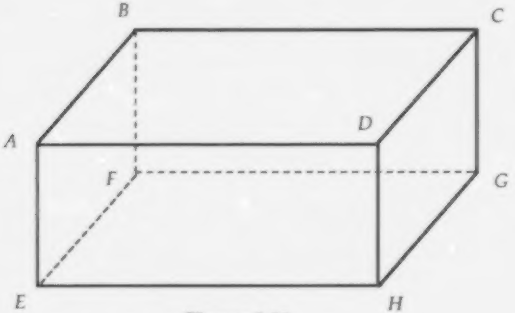


Figure 2.11

The solution to this problem demonstrates the role that imagination as well as insight plays in problem solving. Open the box along several edges to flatten it out. The illustration below is one possibility. $\overline{BH}$ is the shortest path between B and H.

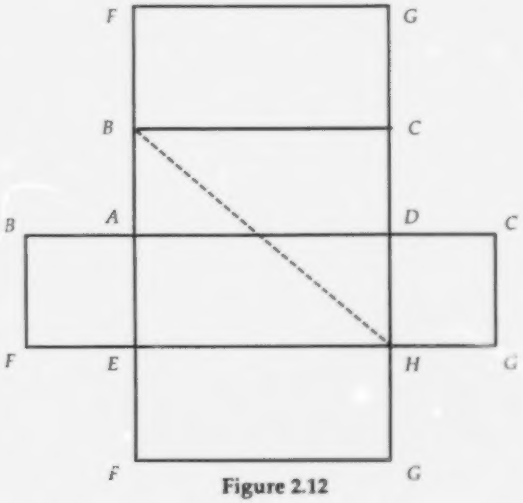


Figure 2.12

Asking Key Questions

The solution to some problems hinges on the solver asking the right kind of question, or "key question," to open the door to a clever solution to the given problem. Not only "crazy people" talk to themselves—so do people trying to solve a mathematics problem. Consider the following problem:

> Suppose you have 7 pairs of blue socks all alike, and 7 pairs of red socks all alike, scrambled in a drawer. How many socks must be taken out at the same time (in the dark) to be certain of getting a matching pair?

COMB

As you begin to tackle this problem, you must ask the right type of key question. To be certain of getting a matching pair, we must assume the worst possible luck. We therefore ask, What is the greatest number of socks, each of a different color, which can be taken from the drawer? Since there are only two colors, the answer is two. Thus the third sock must match one of the first two, so we would have to take three socks from the drawer to be sure of having a matched pair (regardless of color). (See Chapter 4, p. 123)

We shall consider a second problem which also requires asking the right key question.

> An apartment building has 20 apartments and 20 letterboxes. How many letters would a mailman have to deliver to be certain of putting at least 3 letters in one box?

COMB

To solve this problem we search for the right question. We ask, How many letters would be needed to give each letterbox 2 letters? Although these 40 letters could be distributed so that one box has more than 3 letters, we must assume the worst possible luck, namely, that each letterbox has exactly 2 letters. Then the 41st letter would increase the contents of one of these letterboxes to 3 letters. Therefore, 41 letters are needed to be certain that at least 3 letters are put in one box.

To further reinforce the notion of asking the key question, we shall consider the following problem:

COMB

In a drawer you have 6 red ribbons, 7 green ribbons, 4 blue ribbons, and 9 yellow ribbons all the same length and of the same material. You cannot see the colors of the ribbons because the room is dark. How many would you have to take out of the drawer in order to be certain of getting at least one of each color?

Success with this problem once again depends on your asking the right questions. Assuming the worst possible luck, what is the greatest number of ribbons which can be taken from the drawer to have only 3 different colors? We would take all the ribbons of the most plentiful colors: 6 red ribbons, 7 green ribbons, 9 yellow ribbons.

We can have 22 ribbons and still only have 3 of the 4 colors represented. However, the 23rd ribbon must be of a different color! Thus, after 23 picks, we must have at least one of each color. (This is not to say that this could not have been achieved on the first four picks, but it was not to be assumed.)

A geometric problem also can have a solution which hinges on asking the right key question.

GEN

⑥

Consider the following triangular arrangement of checkers. What is the least number of checkers that would have to be moved in order to reverse the direction of the triangle?

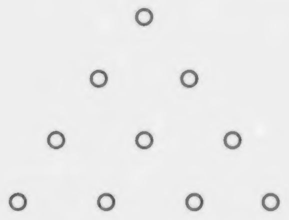


Figure 2.13

Here, the key question is, "How many checkers would remain unmoved if the direction were to reverse?" or "Which checkers are positioned independent of the triangle's direction?"

The following 7 dark checkers can remain unmoved, while the 3 light ones would move as indicated.

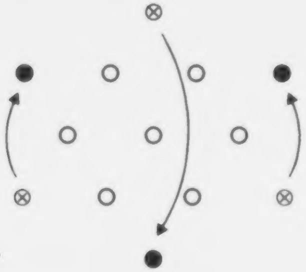


Figure 2.14

The key to the solution to some problems lies in the question one asks (oneself) in the process of looking for a solution. Asking the right question alone may not necessarily yield the solution immediately. Other techniques could also be required. However, stepping back from the problem, reconsidering it from a different perspective, and then asking the right (often unusual) key question, may prove to be the key to a solution. Only practice will ensure success with this notion!

Seeking Complements ③

Notice that sometimes the "key question" is one that asks the "opposite" or complement of what we seek. This is a helpful approach but does not always lead to the right question. The next example will demonstrate how the key question again asks for the opposite of what is being sought.

> To conserve the contents of a 16 oz. bottle of wine, an alcoholic adopts the following procedure. On the first day, he drinks 1 oz. of wine and then refills the bottle with water. On the second day, he drinks 2 oz. of the mixture and then refills the bottle with water. On the third day, he drinks 3 oz. of the mixture and again refills the bottle with water. The procedure is continued for succeeding days until the bottle is empty. How many ounces of water does he drink?

GEN

It is very easy to get bogged down with a problem like this. Instead of trying to find the water content of each day's mixture, simply look to ask the key question: "How much water has been added to the mixture each day?" This key question asks for the "complementary substance" to the one required.

On the first day 1 oz. of water is added.
On the second day, 2 oz. of water is added.
On the third day, 3 oz. of water is added. ③
On the fifteenth day, 15 oz. of water is added.

Why was there no water added on the sixteenth day?

Therefore, $1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 + 11 + 12 + 13 + 14 + 15 = 120$ ounces of water have been consumed.

An alternate solution would involve asking the question, "How much has he drunk in total?" and then deducting the alcohol (16 oz.). Thus,

$$1 + 2 + 3 + 4 + \ldots + 15 + 16 - 16 = 120.$$

In a similar vein, the following problem's solution becomes simple when the key question is asked.

GEN

A tennis tournament has 36 players. One loss eliminates a player. How many games must be played (or defaulted) to get a single winner?

③

Instead of counting the number of games the winners must play, ask the right question (looking for the complementary component): How many losers are there in the tournament? Since there is only one winner, there must be 35 losers, requiring 35 games for the tournament.

How would the solution change if the problem required two losses to eliminate a player?

The key question to open the door to a problem's solution might be the indicator of the key variable to be found. Such is the case in the following problem.

GEN

ALG

③

Two trains, 200 miles apart, are traveling toward each other on the same track. One travels at 60 mph, while the other travels at 40 mph. A superfly perched on the front of the slower train flies toward the faster train at 240 mph, and when it reaches the train, it instantly turns around and flies back toward the slower train (again at 240 mph). When it reaches the slower train it again turns around and instantly flies towards the faster train. This continues until the two trains crash and smash the fly. What was the total distance that the fly traveled?

To try to find the distance, directly, that the superfly traveled altogether would be complicated, at best. The key question here is, What is the time required for the fly's travel? Since we already know the speed of the fly, the distance traveled would simply be the product of the time and the speed. The fly's time of travel is actually the time required for the two trains to collide.

Their times of travel are equal, say t:

Slower train's distance is $40t$.

Faster train's distance is $60t$.

The total distance traveled by the two trains is $40t + 60t = 200$ miles, and $t =$ 2 hours.

Therefore the collision occurs in 2 hours.

The fly traveled for 2 hours at 240 mph and therefore covered a total distance of 480 miles.

The last two problems show solutions that made use of asking key questions, which (at first glance) do not directly aim at the solution sought. The questions are directed at finding some other information within the problem situation. With this additional information, we illuminate the problem further so as to render a solution to the original problem quite simply. We see from these examples that the solution to a problem can sometimes be best found by a somewhat circuitous route!

In this chapter we have considered problems that have lent themselves to some rather unusual solution methods. The problems were specially selected to show these problem-solving strategies in a dramatic fashion. Imitation is often seen to be a useful tool of problem solving. We sometimes recall problem-solving procedures without actually remembering the context in which they were used. Students who are shown these examples (and others like them) should get a "feel" for these methods and thereby begin to develop a fine reserve of problem-solving strategies to use in their study of mathematics as well as in everyday life. Often, problem-solving strategies in the mathematics context can be interchangeable with their use in everyday life. Students should begin to see problem solving not just a means to an end, but rather an end in itself.

3

Interest Grabbers

Exciting Motivational Problems With Punch and Personality

STEVEN R. CONRAD

Introduction

There's nothing quite as stimulating to a young student's mind as a good problem. Plato's *Dialogues* taught us that the best follow-ups to a good question are more good questions. Today, effective teachers frequently capitalize on this advice by using well-chosen problems as a means of instruction. The very best problems are easy to state, easy to understand, easy to remember, and effective at enabling a worthwhile concept to capture our interest. That's a tall order, but it's one that pays enormous dividends. There's a caveat here: Since the heart of a problem often lies in its solution, the value of a problem usually depends on the background and ability of the problem solver.

Problem solving is not a spectator sport. The best problems involve us as participants. We must work at extending our knowledge, we must spend time asking ourselves questions not directly asked of us, and we must reflect on the significance of these activities. In learning a foreign language, "immersion" is the most successful approach. In much the same way, the route to success in mathematical problem solving is immersion—in problem-solving activities.

George Polya taught us that certain techniques are in common use by successful problem solvers. The very first section of Polya's classic *How to Solve It* (Princeton, NJ: Princeton University Press, 1945) is his list of four steps for solving problems. Polya's second step, "Devising a Plan," is the longest step. It concentrates on finding a "connection between the data and the unknown," and it begins with the questions "Have you seen it before?" and "Do you know a related problem?" Indeed, the more experienced we are at problem solving, the better we are able to solve new and more difficult problems. Successful problem solvers carry an arsenal of weapons; they have already seen or solved a wide variety of problems and have thereby learned a wide variety of techniques.

If you want to become a successful problem solver, the most important step is to immerse yourself in solving problems. Athletes train for many years—and continue to train throughout their careers. In much the same way, the success of students who participate in math contests (mathletes) improves when time is spent solving problems.[1] I am reminded of a story told to me some years ago by a friend. When my friend, who had prided himself on his mathematical ability, was given his first homework assignment in math class at college, he whizzed through all the problems save one. Not wanting to hand in an incomplete assignment the very first day, he was determined to solve the remaining problem. Day turned to night, and then back to day. He'd spent all night on the problem and had not solved it. The next day, in class, the professor asked if anyone had any question they wished to ask about any homework problem. My friend raised his hand and asked, "How do you begin problem 6 on page 3?" The professor opened his book, turned to the page, and read the problem to himself. "Problem 6 on page 3?" he inquired. "Cannot be solved," intoned the professor. My friend was visibly shaken. "What do you mean, 'Cannot be solved'? I spent all night working on that problem." The professor looked at him kindly and said, "Well then, you must have learned a lot!"

A Potpourri of Problems and a Medley of Methods

Study the Role of Each Variable

Often, knowing what an equation says about a particular variable—or solving for that variable in terms of the others—gives you new insights. Most people find the following example quite easy to understand but much more difficult to solve:

What are all ordered pairs of unequal positive integers (a,b), with $a > b$, which satisfy $\frac{1}{9} = \frac{1}{a} + \frac{1}{b}$? ALG

Solution 1

What does the equation say about bounds on the possible values of a and b? Since $a > 0$ and $b > 0$, neither can equal 9. Let's prove this. Suppose $b \le 9$. Then $1/b \ge 1/9$ and $1/a \le 0$. Since $a > 0$, this is impossible. Since $a > b$, the values of a and b ⑤ must satisfy $a > b > 9$. Let $a = 9 + x$, and let $b = 9 + y$, with $x > y > 0$. We get $1/9 = 1/(9 + x) + 1/(9 + y)$. Clearing fractions and simplifying the result, we get $xy = 81$. Since $x > y$, $(x,y) = (81, 1)$ or $(27, 3)$. Finally, $(a,b) = (9 + x, 9 + y) = (90, 10)$ or $(36, 12)$.

The example is of a general form, as is its solution. We can summarize by saying:

$$\text{All solutions of } \frac{1}{n} = \frac{1}{n+x} + \frac{1}{n+y} \text{ satisfy } xy = n^2.$$ ⑱

Solution 2

Let's concentrate on the possible values of b, the smaller of the two variables. As in Solution 1, $b > 9$. If both a and b were bigger than 18, the sum of their ⑤ reciprocals would be less than 1/9. If both a and b were smaller than 18, the sum of their reciprocals would be greater than 1/9. Therefore, one is smaller than 18 ⑪ and the other is bigger than 18. Since $a > b$, $a > 18$ and $b < 18$. Concentrating on b and combining requirements, $9 < b < 18$. The possible values for b are 10, 11, . . . , 17. Let's try all 8 possibilities and take note when a is integral. We see that if $b = 10$, then $a = 90$; and if $b = 12$, then $a = 36$.

Solve for One Variable in Terms of the Other

Another way to concentrate on the value of one variable—and let the value of the other follow—is to solve for that variable. Let's use this approach for the previous problem. Clearing fractions, $ab = 9a + 9b$. Solving for b (or a, if you prefer), you get $b = 9a/(a - 9)$. Continue dividing to get $b = 9 + 81/(a - 9)$. In order for b to be an integer, $a - 9$ must be a divisor of 81. Since $a > b$, $a - 9 = 81$ or $a - 9 = 27$ (using any other divisor of 81 would make $a < b$). The two solutions readily follow.

At first glance, the next problem looks nothing like the previous one. The problem statements appear very different—but we have another opportunity to concentrate on one variable. An equation whose solutions must be integers is called a *Diophantine equation*. Try to solve this Diophantine equation:

ALG The roots of $Ax^2 + Bx + C = 0$ are A and B, where A, B, C are nonzero integers. Find A, B, C.

Solution

Before trying to focus on one variable, use any information readily available. The roots are A and B and their sum is $-B/A$, so we have an equation in A and B: $A + B = -B/A$.

③ Solving for B, we get $B = \dfrac{-A^2}{A+1}$

⑯ Dividing, $B = -A + 1 - \dfrac{1}{A+1}$

In order for B to be integral, $A + 1$ must be 1 or -1. Since $A \neq 0$, $A + 1 = -1$, and $A = -2$. From here, $B = 4$. Since B is a root and $B = 4$, we know that 4 is a root of $-2x^2 + 4x + C = 0$. Substituting $x = 4$ into the previous equation, we get $C = 16$.

Additional Problems to Consider

Solve each of the following equations independently. Solutions should be in integers (not just positives):

$$\frac{1}{9} = \frac{1}{a} + \frac{1}{b};$$

$$\frac{2}{9} = \frac{1}{c} + \frac{1}{d}; \text{ and}$$

$$\frac{2}{37} = \frac{1}{e} + \frac{1}{f}.$$

Divide and Conquer

In the previous example, we simplified $b = \dfrac{9a}{a-9}$ to $b = 9 + \dfrac{81}{a-9}$ by dividing the denominator into the numerator. Polynomial division can sometimes offer a dramatic simplification when other insights are not forthcoming. Consider the following problem:

What is the least positive integral value of n for which $\frac{n-12}{5n+23}$ is a non-zero reducible fraction? NUMB

Solution

A non-zero fraction is reducible if and only if its reciprocal is also reducible. ③ By division, ⑯

$$\frac{5n+23}{n-12} = 5 + \frac{83}{n-12}.$$

Now, $83/(n-12)$ is reducible if and only if 83 and $n - 12$ have a common factor greater than 1. Since 83 is a prime, $n - 12 = 83$ gives the least solution, which is $n = 95$.

This solution has a minor trouble spot: That numerator gets in the way of a proper understanding of the relationship between the numerator and the denominator. Since we'd prefer both parts of the fraction to be expressible in as simple a manner as possible, we could let $x = n - 12$. Then the original fraction becomes ④ $x/(5x + 83)$. Take the reciprocal now, and the solution given previously becomes more apparent. In fact, one might be able to find our solution without taking the reciprocal. So frequently do we find that a problem is significantly simplified by a suitable substitution, that the method of substitution is, itself, one of the most important methods of problem solving.

An Additional Problem to Consider

Prove that the fraction $\frac{21n+4}{14n+3}$ is irreducible for every positive integer n. NUMB (This was problem 1 in the First International Mathematics Olympiad, in 1959. Of course, it is difficult!)

The KISS Principle: Keep It Simple—Substitute!

Every problem solver has been confounded by an overly complex problem that became simpler following a certain substitution. Patterns suggest the substitutions.

ALG

Factor completely: $(x+1)(x+2)(x+3)(x+4)-3$.

Solution

② Regroup to get $[(x+1)(x+4)][(x+2)(x+3)]-3$, or $[x^2+5x+4]\,[x^2+5x+6]-3$. Now we can see where to substitute! Let $y=x^2+5x$ to get $[y+4][y+6]-3$. Expanding, we get $y^2+10y+21=(y+7)(y+3)=(x^2+5x+7)(x^2+5x+3)$.

An Additional Problem to Consider

ALG

Solve for x: $(x-1)(x-2)(x-3)(x-4)=48$.

When in Doubt, Generalize

Be careful not to let the previous example mislead you. Often, a good substitution makes a problem simpler but not yet easily accessible. Sometimes, the difficulty can be resolved by considering the general case instead, because in some situations, the general case is easier to handle than a specific case!

ALG

What are all values of x that satisfy

$$\frac{2x^3-3x^2+x+1}{2x^3-3x^2-x-1}=\frac{3x^3-x^2+5x+13}{3x^3-x^2-5x+13}?$$

Solution

The thought of clearing fractions and solving the resulting equation should be abandoned with all due speed! Instead, let's substitute where obvious and take a look at the resulting equation. This does introduce new variables, so things may ⑱ not improve (but it's better to have tried and failed than never to have tried at all). ② Let $a=2x^3-3x^2$, $b=x+1$, $c=3x^3-x^2$, and $d=5x-13$. Substituting, $\dfrac{a+b}{a-b}=\dfrac{c+d}{c-d}$.

Wouldn't those numerators look simpler if instead we could add the denominator to the numerator in each fraction, or subtract the denominator from the numerator instead? Well, we can do both! Let me explain: Adding 1 to each fraction, we get $2a/(a-b) = 2c/(c-d)$. Subtracting 1 from each fraction, we get $2b/(a-b) = 2d/(c-d)$. This is better than the original equation, but still not as good as we can get. Since the equations $2a/(a-b) = 2c/(c-d)$ and $2b/(a-b) = 2d/(c-d)$ have identical denominators, we can divide the first by the second, and then simplify the result, to get $a/b = c/d$. We did it! We showed that whenever $(a+b)/(a-b) = (c+d)/(c-d)$, it then follows that $a/b = c/d$ (which is so much simpler to solve). Now, substitute back to get

$$\frac{2x^3-3x^2}{x+1} = \frac{3x^3-x^2}{5x-13}, \text{ or } \frac{x^2(2x-3)}{x+1} = \frac{x^2(3x-1)}{5x-13}.$$

If $x = 0$, the equation is satisfied. If $x \neq 0$, we can divide both sides by x^2 to get the much more easily solvable $(2x-3)/(x+1) = (3x-1)/(5x-13)$.

Clearing fractions, $7x^2 - 43x + 40 = (7x-5)(x-8) = 0$. The solutions are 0, 5/7, 8.

Additional Problems to Consider

1. What are all values of x which satisfy ALG ②

$$\frac{3x^4+x^2-2x-3}{3x^4-x^2+2x+3} = \frac{5x^4+2x^2-7x+3}{5x^4-2x^2+7x-3}?$$

2. What are all values of x which satisfy ALG ②

$$\frac{\sqrt{x+1}+\sqrt{x-1}}{\sqrt{x+1}-\sqrt{x-1}} = \frac{4x-1}{2}?$$

It's Based on How You Look at It

If asked how to drive from New York, in the northeastern United States, to a small town in the southwest, one would not begin with a map of New York: There are just too many routes that leave New York. Where would you begin? Begin at your destination and try to find the best way back to New York! ①

In the Unlucky Lottery, all the prizes are powers of \$13 (\$1, \$13, \$169, etc.), and the total prize money to be given away is \$1 million. What is the least possible number of prizes in the Unlucky Lottery? ARITH

Solution 1

① ⑨ ⑪ The prize amounts are $1, $13, $169, $2,197, $28,561, and $371,293. To give the least *number* of prizes, start with the largest prize: award as many of these as possible before giving out smaller prizes. There can be at most 2 prizes of $371,293; and then $257,414 would remain to be awarded. We could then give away 9 prizes of $28,561. This would leave $365, which could be most efficiently given away as 2 prizes of $169, 2 prizes of $13, and 1 prize of $1—for a total of 16 prizes.

Is there a way to do all this more simply—or at least to describe the procedure more easily? Yes! Use thirteen as the base of numeration, as in Solution 2:

② ③ ⑨ ⑪ *Solution 2*

In base thirteen, the number 1 million is 290221, meaning there are 2 prizes of $371,293, 9 prizes of $28,561, 2 prizes of $169, 2 prizes of $13, and 1 prize of $1. The least possible number of prizes is 16.

An Additional Problem to Consider

ARITH ② How can $1 million in prizes be given away in the Lucky Lottery, if all prizes are powers of $11, and if no more than 7 people can receive the same prize?

Consider a Related Problem

Though there are several ways to sum the series in the next problem, none of them is obvious. Give yourself time to puzzle over the following problem for quite a while before looking at the solution. The best approach to solving this problem is to first investigate a very similar series that you can sum. You'll enjoy this problem.

ALG What is the sum of n terms of the series 7 + 77 + 777 + 7777 + . . . , where the k^{th} term is a k-digit number, each of whose digits is 7?

Solution

④ Let's sum n terms of 9 + 99 + 999 + . . . , instead. We can rewrite this series as (10 – 1) + (100 – 1) + (1,000 – 1) + . . . The sum of n addends will be 10 + 100 + 1,000 +

10,000 + . . . $-n$. The powers of 10 form an n–term geometric series whose sum is $\frac{10^{n+1}-10}{9}$ so the sum of n terms of 9 + 99 + 999 + . . . is $\frac{10^{n+1}-10}{9} - n$. But what about the original problem?

Since 7 is 7/9 of 9, the sum of n terms of the original series is

$$\frac{7}{9}\left(\frac{10^{n+1}-10}{9} - n\right) = \frac{7}{81}(10^{n+1} - 9n - 10).$$

It is interesting to note that, for n ≤ 9, the sum is 7(123456 . . .), where the last digit used inside the parentheses is the value of n. (When n ≥ 10, this method does not work.) ③ ⑰

Additional Problems to Consider

1. What is the sum of the infinite series $1/3 + 2/3^2 + 3/3^3 + \ldots + \frac{n}{3^n} + \ldots$? ALG
2. What is the value of $2^{\frac{1}{2}} \times 4^{\frac{1}{4}} \times 8^{\frac{1}{8}} \times 16^{\frac{1}{16}} \times \ \ldots (2^n)^{\frac{1}{2n}}$? ALG
3. What is the sum of the infinite series $\sum_{n=1}^{\infty} \frac{n^2}{3^{n-1}}$? ALG

Can You Count? (The Inclusion-Exclusion Principle)

If you had 2 coins made of gold, and you had 2 coins over 20 years old, you could have 2, 3, or 4 coins altogether—the total depends on how many of your gold coins are over 20 years old. Any time you count, remember to account for the possibility of duplicates when objects can be categorized in more than one way.

There's an analogous result in set theory: For two sets, if $n(S)$ is the number of elements in set S, then $n(A \cup B) = n(A) + n(B) - n(A \cap B)$. The next problem arose from the observation that, although $\sqrt{9} = \sqrt{4} + \sqrt{1}$, and $\sqrt{18} = \sqrt{9 \times 2} = 3\sqrt{2} = 2\sqrt{2} + \sqrt{2} = \sqrt{8} + \sqrt{2}$, the number $\sqrt{10}$ *cannot* be written as a sum of square roots of integers. ⑪

If a and b are positive integers, the equation $\sqrt{10} = \sqrt{a} + \sqrt{b}$ has no solutions. For how many positive integral values of $x \leq 1{,}000$ does the equation $\sqrt{x} = \sqrt{a} + \sqrt{b}$ have at least one solution in which both a and b are positive integers? COMB

Solution

⑪ The equation $\sqrt{x} = \sqrt{a} + \sqrt{b}$ has a solution in positive integers if and only if x contains a perfect square factor. (This assertion is proved as a note following this solution.) For example, $\sqrt{9} = \sqrt{4} + \sqrt{1}$, or $\sqrt{18} = \sqrt{9 \times 2} = 3\sqrt{2} = \sqrt{2} + 2\sqrt{2} + \sqrt{2} + \sqrt{8}$. There are 250 values of $x \le 1{,}000$ that contain a factor of 4. Similarly, the number of values of $x \le 1{,}000$ that (respectively) contain a factor of $3^2, 5^2, 7^2, 11^2, 13^2, 17^2, 19^2, 23^2$, and 31^2 is 111, 40, 20, 8, 5, 3, 2, 1, 1, and 1—for a total of 442 values. But some of these values, such as 36, have been counted twice—once as a multiple of 2^2 and once as a multiple of 3^2. We must count the number of such duplications and subtract them from our total. The number of values of $x \le 1{,}000$ that respectively contain a factor of $2^2 3^2, 2^2 5^2, 2^2 7^2, 2^2 11^2, 2^2 13^2, 3^2 5^2$, and $3^2 7^2$ are 27, 10, 5, 2, 1, 4, and 2—a total of 51 such duplications. Finally, there is the case of $2^2 3^2 5^2$. This was counted 3 times in the first group, and was then deleted 3 times in the second group (for $2^2 3^2$, $2^2 5^2$, and $3^2 5^2$). Thus, the answer is $442 - 51 + 1 = 392$.

ALG ⑱ (**Note:** Let's prove that there is a solution in positive integers if and only if x contains a perfect square factor. Since $\sqrt{x} = \sqrt{a} + \sqrt{b}$, $x = a + b + 2\sqrt{ab}$. In order for x to be an integer, ab must be a square. If the prime factorization of a has a prime factor to an odd power, then b must also have that same prime factor, also to an odd power (and *vice versa*). Thus, when a is simplified, say, into the form $u\sqrt{v}$ then b must simplify into the form $w\sqrt{v}$. Therefore, $\sqrt{x} = u\sqrt{v} + w\sqrt{v} = (u + w)\sqrt{v}$, and $\sqrt{x}$ must be of this form. Finally, $x = (u + w)^2 v$, and x is a square or a multiple of a square. We can prove the converse by working backwards.)

Additional Problems to Consider

NUMB 1. How many positive integers less than 1,000 are divisible by neither 5 nor 7?

NUMB 2. For how many integers x between 20,000 and 60,000 is x the square of an integer or the cube of an integer?

NUMB 3. How many integers from 1 to 100 inclusive have no repeated prime factor?

Use the Basic Operations to Simplify Matters

In solving difficult algebraic equations of the types given in contests, the key step often involves combining the equations in some nifty way that makes the solution apparent. Recall that an algorithmic way to solve a two-variable linear system is to multiply each equation by its own multiplier, then add the results. This *linear combination* of the two equations is often all we need in order to solve the system. Though a calculator trivializes the next problem (since the problem's only difficulty is the size of its coefficients), the original wording has not been revised. It is hoped that when those who love mathematics see the symmetry in the equations, they'll eventually seek an "artistic" approach.

With each major curriculum change, we gain something big and lose something small. There's a beauty in some parts of mathematics that is best witnessed with technology. And there's a beauty in other parts that will be lost. Perhaps the next problem will remind us of an era in which we could find much cleverness in parts of arithmetic.

What is the ordered pair of real numbers (x,y) for which

$$123x + 321y = 345$$

ALG

and

$$321x + 123y = 543?$$

Solution

Adding, we get $444x + 444y = 888$; or $x + y = 2$. Subtracting the first of the original equations from the second, we get $198x - 198y = 198$; or $x - y = 1$. Solving $x + y = 2$ and $x - y = 1$, we get

$$(x, y) = \left(\frac{2}{3}, \frac{1}{2}\right).$$

Additional Problems to Consider

1. Solve for (x, y): $713x + 637y = 4{,}164$ and $637x + 713y = 3{,}926$. ALG
2. If each of four numbers is added to the average of the other three, the respective sums are 28, 32, 40, and 44. What are the four numbers? ALG

Watch the Domain

An ability to analyze domains is basic to understanding mathematics. Among high school students, an ability to handle domain considerations properly is a hallmark of mathematical skill. Solving domain problems is an excellent way to review fundamental properties of functions. Though we often give them but brief mention, domain considerations are sometimes the only window through which we can view the essence of a problem. The next example is remarkable in several ways: First, it uses each of the six trigonometric functions and their inverses; second, though no number appears in the problem itself, the answer is numerical; and third, though the problem statement involves neither an equation nor an inequality, the solution involves both. There is no other problem quite like this one—it's a one-problem course in trigonometry!

TRIG

Evaluate the real function:
sin arc tan sec arc csc cot arc cos csc arc cot cos arc sin tan arc sec y.

Solution

⑧ ⑰

If there is an arc sec y, then $|y| \geq 1$. Next, since tan arc sec $y = \pm\sqrt{y^2 - 1}$, we know that cos arc sin $(\pm\sqrt{y^2 - 1}) = \pm\sqrt{2 - y^2}$, so $|y| \leq \sqrt{2}$. We know that the value of csc arc cot $(\pm\sqrt{2 - y^2})$ is $\pm\sqrt{3 - y^2}$. But the cosine of an angle cannot exceed 1 in absolute value, so in arc cos $(\pm\sqrt{3 - y^2})$, $|y|$ must equal $\sqrt{2}$; and arc cos $(\pm 1) = 0$ or π. Then, cot arc cos (± 1) is undefined, and sec arc csc(undefined) is ± 1. Finally, the value of sin arc tan (± 1) is $\pm\sqrt{2}/2$.

Additional Problems to Consider

Though each of the problems below can be done by traditional methods, domain considerations provide a welcome headstart.

TRIG
TRIG
ALG

1. Solve the equation $2\sin x = 5x^2 + 2x + 3$.
2. Solve the equation $\cos^7 x + \sin^4 x = 1$.
3. Solve the equation $x = \sqrt{x^2}$.

Draw a Line from the Center

Years of experience solving math problems has two effects: It ages you (the years, not the experience), and it gets you started with the right strategy more quickly (the experience, not the years). When solving a circle problem, a strategy that often works well is to draw one or more strategic radii. An additional strategy useful in more difficult circle problems is to draw one or more lines through the center, perpendicular to another line. The actual technique varies with the problem at hand, so try your hand at the three problems below:

Squares of side-length 6 are inscribed in sectors of a circle, as shown below. What is the ratio of the area of the square to the area of the sector when it's a semicircle, when it's a quadrant, and when it's an octant?

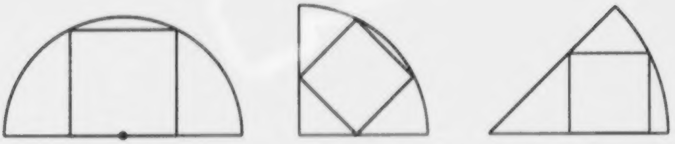


Figure 3.1

Solution

Remember that a sector with radius r and central angle α (in radians) has area $\frac{1}{2}r^2\alpha$. To evaluate r, draw the lines shown and use the Pythagorean theorem.

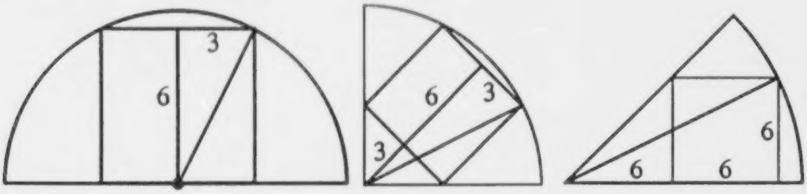


Figure 3.2

An Additional Problem to Consider

In all three problems above, the ratio sought turns out to be $8/5\pi$. Is it always $8/5\pi$? Not at all! Let's consider the case where exactly two of the vertices of the square lie on the arc, as in the first two drawings. By drawing a radius and a center line perpendicular to the far side of the square, show that if the central angle is 2θ (in radians), then the ratio sought is always $4/[(5 + 4\cot\theta + \cot^2\theta)\theta]$.

(*Note:* In the first two diagrams, $\theta = \pi/2$ and $\theta = \pi/4$, respectively.)

Do You Know a Related Theorem?

Construction problems are not as popular as they once were. But constructing figures that satisfy a given condition is really an exercise in correctly remembering and applying the theorems learned in geometry, so constructions are among the best (and also among the most difficult) problems in geometry.

Draw two circles intersecting at two points so that the degree measure of the arc cut off on one circle is twice that of the arc cut off on the other.

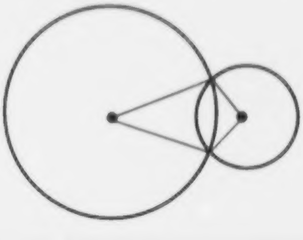


Figure 3.3

Solution

① ⑥ ⑯ Construction problems are often handled by first drawing a picture of the construction as completed, then next identifying theorems that apply to that figure, then finally working backwards to a solution. Here, one theorem stands out: *In a circle, if a central angle and an inscribed angle intercept the same arc, then the central angle is twice the inscribed angle.*

How does that help? We need to draw an inscribed angle in the smaller circle—the same circle in which we already have the central angle—and we'll need to require that the sides of this inscribed angle, together with the sides of the central angle of the larger circle, form a rhombus. Then, since opposite angles of a rhombus are congruent, we're on our way! Working backwards, we now *begin* by drawing a rhombus. Next, draw a circle whose center is vertex V of the rhombus and which passes through two other vertices of the rhombus; and draw a second circle that passes through every vertex of the rhombus *except* vertex V.

An Additional Problem to Consider

⑥ ① GEOM Construct a triangle given the lengths of all the medians. (*Note:* As before, begin by drawing the completed triangle. Then, draw in the medians. To complete the construction, you'll have to construct a parallelogram, given two sides and a diagonal. Which parallelogram? That's the difficult part, but the secret lies in the figure you've already drawn—so study it carefully!)

Know How the Roots Relate to the Coefficients

In every polynomial equation with lead coefficient 1, the other coefficients are functions of the roots. For example, if the cubic equation $x^3 + px^2 + qx + r = 0$ has roots a, b, and c, then the sum of its roots, taken one at a time, is $-p$, so $a + b + c = -p$; the sum of its roots, taken two at a time, is q, so $ab + ac + bc = q$; and the sum of its roots, taken three at a time, is $-r$, so $abc = -r$. Use these facts to solve the next problem:

ALG What are three numbers whose sum is 13, whose product is –165, and the sum of whose squares is 155?

Solution

③ ② Let's reconstruct the equation $x^3 + px^2 + qx + r = 0$. Since $a + b + c = 13$, $p = -13$. Since $abc = -165$, $r = 165$. So far, $x^3 - 13x^2 + qx + 165 = 0$. We still need to know the value of $q = ab + ac + bc$. We know that $ab + ac + bc$ is a sum of second-degree terms. But so is $a^2 + b^2 + c^2$. Let's link these in a single equation. Square both sides of $a + b + c = 13$ to get $a^2 + b^2 + c^2 + 2(ab + ac + bc) = 169$. The sum of the squares of the three numbers is 155, so $155 + 2(ab + ac + bc) = 169$, and $ab + ac + bc = 7$. We've done it: An equation with our three numbers as roots is $x^3 - 13x^2 + 7x + 165 = 0$. Factoring, $(x-5)(x^2-8x-33) = 0$ or $(x-5)(x+3)(x-11) = 0$, and our three numbers are -3, 5, and 11.

Additional Problems to consider

1. For what values of p, q, and r are p, q, and r the roots of $x^3 - px^2 + qx - r = 0$? ALG
2. The roots of $x^3 + px^2 + 3x + r = 0$ form an arithmetic progression whose common difference is 3. What is the sum of the absolute values of the roots? ALG
3. If a, b, and c are different numbers, and if $a^3 + 3a + 14 = 0$, $b^3 + 3b + 14 = 0$, and $c^3 + 3c + 14 = 0$, what is the value of $1/a + 1/b + 1/c$? (**Hint:** Since a, b, and c all satisfy $x^3 + 3x + 14 = 0$, then a, b, and c are its three roots. Also, $1/a + 1/b + 1/c = (bc + ac + ab)/abc$.) ALG
4. For what pair of numbers (p, q) are the roots of $x^2 + px + q = 0$ the squares of the roots of $x^2 + x + 1 = 0$? (**Note:** The answer is *very* surprising!) ALG

Making a List, Checking It Twice

In some problems, no equations appear and no inequalities appear—there seems to be no way to gather all the information into one small package. On such occasions, a list, chart, or table often brings all the data together.

A fuel tank receives a continuous, steady flow of 2,000 liters per hour. The tank experiences a steady rate of fuel usage within each of the 6 consecutive 4-hour periods every day. Every day, usage during these periods is respectively 6,000 13,500, 7,300, 10,000, 8,000, and 3,200 liters. What is the capacity, in liters, of the smallest tank that could ensure there would always be at least 200 liters of fuel in the tank?

ARITH

GEN

Solution

Suppose there are x liters in the tank at the start. Then, during each of the 6 consecutive 4-hour periods, the tank would get $4(2{,}000) = 8{,}000$ liters. Thus, at the end of the six periods, the tank would contain, respectively, $x + 2{,}000$, $x - 3{,}500$, $x - 2{,}800$, $x - 4{,}800$, $x - 4{,}800$, and x liters. In order for the smallest of these, $x - 4{,}800$,

⑥ ⑪

to be at least 200, x must be at least 5,000. But the tank must be able to hold $x + 2{,}000$ liters at the end of the first period. Consequently, the minimum liter capacity of the tank must be 7,000 liters.

Additional Problems to Consider

LOGIC

1. After a period of particularly bad weather, the Parliament of Franistan decreed that there would no longer be any weather on even-numbered calendar dates. During one of the months following that decree, three Mondays had no weather. On what day of the week did the 13th of that month occur?

LOGIC

2. Mrs. Smith has three sons: Stan, Steve, and Stu. One is skiing in Selden, a second is in Setauket, and the third is in Smithtown. One is skating, one is skiing, and one is skeet-shooting. Stan is not in Selden, Stu is not in Setauket, and the son who is skeet-shooting is not in Smithtown. If the skiing enthusiast is not Stu, who is skeet-shooting and where?

Investigate All the Possibilities

If $a^b = 1$, *many* people believe that $a = 1$ or $b = 0$. But there is a third possibility. Armed with this knowledge, solve the next problem:

ALG

What are all values of x for which $(x^2 - 5x + 5)^{x^2 - 9x + 20} = 1$?

Solution

⑪

The left side can equal 1 in only 3 cases: first, the base = 1; second, the exponent = 0 and the base $\neq 0$; and third, the base = -1 and the exponent is even. Let's consider each case separately.

Case I. [base = 1]
In this case, $x^2 - 5x + 5 = 1$, so $x^2 - 5x + 4 = (x - 4)(x - 1) = 0$ and $x = 1$ or $x = 4$.

Case II. [exponent = 0, base $\neq 0$]
In this case, $x^2 - 9x + 20 = (x - 5)(x - 4) = 0$; so $x = 4$ or $x = 5$. Both roots check.

Case III. [base = -1, exponent even]
In this case, $x^2 - 5x + 5 = -1$, or $x^2 - 5x + 6 = (x - 3)(x - 2) = 0$. Solving, $x = 2$ or $x = 3$ (in both cases, the exponent is an even integer). Therefore, the values of x are 1, 2, 3, 4, 5.

Additional Problems to Consider

1. Solve $4^{x^3+5x^2-6x} = 1$. (There are three solutions.)
2. Solve $x^{x^2-7x+12} = 0$. (There are four solutions.)
3. Solve $x^{(x+1)^2} = x^{16}$. (There are five solutions.)
4. If $a \uparrow b$ means a^b, what are all values of x for which $x \uparrow (x \uparrow x) = (x \uparrow x) \uparrow x$? (There are three solutions.)
5. What are the three ordered pairs of integers (x, y) which satisfy the simultaneous system $x^{x+y} = y^4$, $y^{x+y} = x$?
6. Solve: $\sqrt{x^{x+1}} = x^{\sqrt{x+1}}$. (There are four solutions.)

ALG

Use an Indirect Proof

To prove a statement false—to disprove a statement—it is often best to try an indirect proof. To disprove proposition *P* indirectly, begin by assuming that *P* is true, and then show that such an assumption leads to a contradiction.

Determine whether or not it is possible to draw a straight line that intersects all sides of a polygon of 999 sides.

COMB

Solution

If such a line existed, then each time it crosses a side of the polygon, one vertex of the polygon would fall on one side of the line, and one would fall on the other side of the line. But, for a polygon of 999 sides, it is not possible to have an equal number of vertices on both sides of the line.

Six circular regions in the plane have the property that none contains the center of any other. Prove that these regions have no point in common.

GEOM

Solution

Assume, to the contrary, that the circular regions do have at least one point in common. Now draw six segments so that each joins this common point to the center of a different one of the six circles. There are only six segments, and the sum of the measures of the six angles about the common point is 360°, so there must be

⑤ ⑨ ⑬ ⑯ at least two segments for which it is true that the angle between them is less than or equal to 60°. Let the lengths of these two segments be *a* and *b*, with $a \leq b$. Now draw a circle, centered at the common point, whose radius-length is *b*. This circle must contain the center of the circle associated with segment *a*. But this center must also be included within the circle associated with segment *b*, which contradicts our original assumption.

Additional Problems to Consider

ARITH 1. Prove that the number 0.12345678910111213 . . ., formed by writing down all the positive integers, in order, after the decimal point, is *not* periodic.

ARITH 2. A 1,000-digit number contains, in some order, 333 ones and 777 zeroes. Can such a number ever be the square of an integer? (Hint: Consider divisibility by 3 and by 9.)

TRIG 3. Prove that the function $f(x) = \cos\sqrt{x}$ is not periodic.

Miscellaneous Examples for Further Consideration

GEN ② 1. Find the maximum and the minimum number of Friday-the-13th's that can occur in a calendar year. Consider both leap and nonleap years.

GEOM ⑨ ⑱ 2. A square has a side of length 2. The midpoint of each side is joined to the two opposite vertices. What is the area of the octagon at the center of the square? Show that, if "square" is replaced by "parallelogram," the area of the octagon is 1/6 the area of the parallelogram.

ALG ⑨ 3. If $a + b + c = 1$, $a^2 + b^2 + c^2 = 2$, and $a^3 + b^3 + c^3 = 3$, show that $a^4 + b^4 + c^4 = 25/6$.

GEOM ⑧ 4. The opposite vertices of an $x \times y$ rectangle are brought into coincidence, and the rectangle is then flattened out to form a crease. Show that the length of the crease is $\left(\frac{y}{x}\right)\sqrt{x^2 + y^2}$.

GEOM ⑥ 5. Which fits better: a round peg in a square hole or a square peg in a round hole?

GEOM ⑧ 6. Prove that no regular polygon with more sides than a square can be inscribed in an ellipse.

ALG ⑱ 7. What are all values of *x* for which $(16x^2 - 9)^3 + (9x^2 - 16)^3 = (25x^2 - 25)^3$?

ALG 8. If *x* is a real number, what is the minimum value of $(x + 8)(x - 8)(x + 6)(x - 6)$?

Reprise

Is it worth it? If you're "doing quite well, thank you," is it worth investing the time needed to introduce techniques and technologies that may be unfamiliar to you but that may improve problem investigation and problem solving? If you are comfortable in the classroom, why try anything new?

First and foremost, it helps us to teach better: There are always some topics we've "got to get going" a little better. It's a matter of the old maxim "If you always do what you've always done, you'll always get what you've always gotten."

Hard work at teaching pays off in the end: When you're alert to new ideas, your students will be alert to you.

Note

1. The best way to understand a problem's solution is to write a new problem based on the concept learned in solving the old one. Some of the finest problems began this way. A few personal favorites even had their inspiration in routine textbook problems. The final result need not look anything like the inspirational model—it is the idea that matters most. What follows are a variety of solution methods and problems which illustrate these methods. Most of these problems were written expressly for Mathematics League competitions. For others, the sources are no longer remembered. After you've worked through a problem, try to write a new one on your own. For additional ideas, you may want to read the problem books coauthored by Dan Flegler and Steven Conrad. These six books (and the regional and national—United States and Canada—contests on which they are based) are available from Math League Press, Box 1090, Manhasset, NY 11030. The problems used in this chapter are generally among the most difficult of the problems you'd ever see on a math contest. Have fun—and learn a lot!

4

Check the Answer, Please!

MARIO SALVADORI

All of us who teach mathematics and use it in our daily activities often solve a problem by the intuition engendered in us by familiarity with our subject. Yet, once in a while, upon deeper reflection, we begin to suspect that our intuition may be misleading us. We then check the solution with pencil and paper, a calculator, or a computer and realize, to our surprise, that we had been wrong. (If this happens to us professionals, how much more often is it bound to happen to our students?)

This experience is common not just to mathematicians but to all scientists and technologists. In my professional activity as an engineering designer, I myself have experienced such doubts and the need to "check the answer." As a forensic engineer, I have encountered cases where *not* "checking the answer" had led the inexperienced or careless practitioner to tragic consequences. A famous case comes to mind where an apparently "minor suggestion" by the contractor, unchecked by the designing engineers, led to the collapse of two walkways, or pedestrian bridges, connecting two sections of a hotel building (in Kansas City, Missouri) at

---

AUTHOR'S NOTE: All of the engineering-design failures mentioned in this chapter are taken from *Why Buildings Fall Down*, by Matthys Levy and Mario Salvadori (New York: Norton, 1992).

the second- and fourth-floor levels. The collapse of the walkways within its main hall killed 114 dancers and maimed almost 200 more people.

To exemplify this kind of occurrence, I will outline the solution of a few problems in mathematical probability (that purest and most useful chapter of mathematics), which on the face of it have "obvious" but incorrect solutions, together with a few additional examples of engineering-design failures due to the lack of careful checking of design drawings or physical hypotheses.

PROB

I suggest that some, if not most, problems in probability are approached incorrectly by our students because of a lack of clear understanding of the so-called fundamental theorem of probability theory, a hypothesis that can be easily demonstrated by its application to a classical problem.

We all know that if we flip an ideally balanced penny, it has the same chance of showing heads as tails. Since, moreover, it can only show either one or the other, the probabilities of heads and of tails must both be ½, because their sum must be 1, the probability of certainty.

It should also be obvious that, since no *preceding* throw can influence a *subsequent* throw (i.e., since the throws, if honest, are independent of each other), the probability of getting one more head after having thrown 20 heads is still ½. Why, then, do most people feel that it is not so and bet tails after getting 20 heads? It is because they confuse the answers to two different questions. First, we ask what the probability is of heads on one particular throw; since the penny is unaware of what has happened before, the chance of heads on the 21st throw is ½, as it was for every preceding throw. In the second question, we involve the previous throws in the count by asking what the probability is of getting 21 heads in a row; of course, we obtain a totally different answer.

For example, if we ask what the probability is of getting two heads in a row, we may get the following combinations of heads (H) and tails (T):

HH, HT, TH, TT.

The probability of HH is one out of four, or ¼—that is, ½ × ½, the product of the probability of ½ for heads in each throw. Similarly, taking into account the preceding throws, the chance of 21 heads in a row have the (extremely small) probability of $(\frac{1}{2})^{21}$, or 1 chance in 2,097,152. Thus, the layman's feeling that the 21st throw of heads has a minimal chance when it is preceded by 20 heads is correct, even if, in fact, the 21st throw itself has the same ½ probability as any other.

We can now state the fundamental theorem of probability: *The probability that the frequency of a sequence of occurrences will approach the theoretical probability of that sequence increases with the number of occurrences.*

In this statement, the frequency of an occurrence is how often it actually happens in a sequence of independent occurrences, and what increases is only the probability that the frequency should approach the theoretical value of the probability. The theorem explains that when we throw the penny the 21st time after getting 20 heads, there is no guarantee that we will *not* get one more head, only guarantees that the probability of getting increasingly many heads will become less and less. With this preliminary (and much needed) clarification, we can now take *chances* knowing what their probabilities are.

PROB

In Chapter 2 (p. 95) of this book a problem is ingenuously solved that *is* a probability problem but leads to the wrong solution if its statement is not carefully read (as all statements should be).

GEN

> Suppose you have 7 pairs of blue socks all alike and 7 pairs of red socks all alike, scrambled in a drawer. How many socks must be taken out at the same time (in the dark) to be certain of getting a matching pair [of *any one* of the two colors]? (Emphasis added)

Treating the problem as one in probability, we realize that if we extract 3 socks at the same time, we may obtain one of the following six combinations:

⑪

BRR, BBR, BBB, RBB, RRB, RRR.

All of these combinations satisfy our demand for the red or the blue socks so that the probability of getting a matching pair is a certainty (i.e., the probability equals 1).

But if we misread the statement and solve the slightly different problem that asks how many pairs of socks must we take out of the drawer to have the *certainty* of getting two socks *of the same color?* (say blue), we may be unlucky enough to take out first as many as 14 red socks in a row before getting 2 blue socks—the worst case, with the tiny probability of $\left(\frac{1}{2}\right)^{14}$, or 1 in 16,384. Fortunately, the fundamental theorem reassures us that the probability of being that unlucky is very small indeed.

Let me mention here another engineering catastrophe due to a lack of careful reading of their own drawings by the designing engineers of an important structure.

The 300 ft. × 360 ft. roof of the Hartford (Connecticut) Civic Center arena consisted of a "space frame" of almost 5,000 interconnected steel bars supported on four pylons of concrete—a structure famous worldwide because it was entirely built on the ground and then lifted hydraulically to the top of the pylons 100 feet above it. The designing engineers had made sure that the highly compressed steel bars of the roof's interior sections would not "buckle" (collapse in bending), a most dangerous phenomenon that occurs suddenly without any warning. They did not realize, however, that the bars along the perimeter of the roof were not prevented from buckling outward—they did not make sure that these 50 bars could support the loads on them. Five years after its erection the roof collapsed on a winter night (at 5 a.m.), dropping to the ground 15,000 tons of steel.

PROB Let us now consider the well-known "birthday problem," which exemplifies the potential confusion that can arise from checking for the reasonableness of an answer to a question. The solutions here are anything but reasonable, but they are nevertheless correct.

There are 30 students in your class. What is the probability B that two of them have the same birthday? (Assume a year of 365 days.)

③ The simplest way of determining the probability B is to "ask the right question" and determine first the probability p of no two students in the class having the same birthday. Choose any one student. His or her probability is 1, or 365/365, since he or she must have been born on some day of the year. The next student has 1 less way out of 365 of having a birthday different from that of the first; his or her probability is (365 – 1)/365. The next has a probability (365 – 2)/365 of having a birthday different from the first 2; the next, a probability (365 – 3)/365 of having a birthday different from the first 3, and so on. The combined probability of all 30 students *not* having the same birthday is thus given by the product of these 30 probabilities:

$$p = (365/365) \times [(365-1)/365] \times [(365-2)/365] \times \ldots$$
$$[(365-28)/365] \times [(365-29)/365].$$

Since the probability B that two students have the same birthday added to the probability p that two students do not have the same birthday is a certainty, they must add up to 1, and

$$B = 1 - p.$$

With a little patience, p can be evaluated on a calculator and is found to be equal to 0.294, or approximately ⅓. Hence the probability B is $1 - 0.294 = 0.706$. The *odds* of two students having the same birthday is the ratio of $B/p = 0.706/0.294 = 2.40$; "amazingly" enough, the odds are greater than 2 to 1 that 2 students among 30 have the same birthday.

This and other interesting probabilities for the birthday problem can be easily obtained by means of the following elementary BASIC computer program, run in double precision on any computer and setting the last number in line 3 equal to the number of people in the group in question.

BASIC COMPUTER PROGRAM FOR THE EVALUATION OF B

```
DEFDBL A - Z
p = 1
FOR i = 2 TO 30
p = p * (365 - i + 1) / 365
NEXT i
B = 1 - p
PRINT ; B
```

The *B* probability for 30 students given above was evaluated by this program, which shows that for a party of 55 people $B = 0.986$—that is, almost a certainty—and for a party of 100 people, $B = 0.9999997$, which to all practical purposes *is* a certainty.

When the Comets, the first commercial jet planes built in Great Britain in 1952, first began to fly internationally in 1954, seven of them crashed inexplicably within 2 years, killing a large number of passengers. It took the Royal Aircraft Establishment a long time and a lot of demanding research to find out the cause of these disasters. Eventually, experiments proved that the fuselage of the Comets had exploded due to minute cracks at the corner of one or more windows. These cracks were the consequence of two physical phenomena: (a) the stresses at the corners of the windows, which were much higher than elsewhere in the fuselage (as they are at the reentrant corners of any opening), and (b) the weakening of a metal element when the stress in it changes *frequently* from tension to compression, from pushing to pulling and back again (you can easily check this phenomenon by bending back and forth the wire of a paper clip: It will break, usually, after 10 to 20 bendings). The designers had not checked to find that both phenomena did occur simultaneously during the long, bumpy flights of the Comets, eventually breaking up the plane after a few thousand miles of flight.

In the following problem, a student shopping at the point of "reasonableness" has no chance of reaching the correct answer. As teachers, we must insist on checks beyond just ensuring reasonableness, though its importance is not to be minimized.

A man has 3 records. The first has songs on both sides (S_1S_2); the second instrumental music on both sides (I_1I_2); and the third songs on one side and instrumental music on the other (S_3I_3). He comes into his living room, puts on one of the records, and hears a song. What is the probability that the other side of the same record also has songs?

PROB

GEN

Since the 3 records have altogether 3 song sides and 3 instrumental sides, all equally likely, your students may think, at first, that the required probability is ½. Some of them may even believe, by analogy with two consecutive heads or tails in the penny-throws problem, that the probability of two "consecutive" songs is ½ × ½ = ¼. But the two sides of the records are *not* independent, and the (equally probable) ways of putting on one of the two records with a song on either one or two sides are 3:

$$S_1/S_2,\ S_2/S_1,\ S_3/I_3,$$

of which the first 2 satisfy the question. Hence, the probability that the man will hear a song on the other side of the record now playing is ⅔. (Neat, but tricky at first.)

Another example of a design failure due to an unusual physical phenomenon requiring careful checking is the collapse in 1979 of the 324 × 360 ft. roof of the magnificent Kemper Arena in Kansas City, Missouri, 6 years after its erection. On an exceptionally rainy night, water started to accumulate on the roof of the arena, bending it slightly into a dished shape. As more and more rain fell, accumulating more and more water on the roof for lack of sufficient spouts, the roof dished more and more until this self-increasing phenomenon, called "ponding," allowed so much water to accumulate on the roof that its weight collapsed it. Fortunately, the arena was not in use that day. (Perhaps because the evaluation of ponding of this complex roof demanded such unusual calculation, or perhaps because it was considered wise not to waste too much money on lawyers, the case was settled in court in 2 days, without involving the engineers' responsibility, despite their obvious failure in designing the roof's water spouts.)

In order to pass time when sitting in a car next to a silent driver, I often play the "car-plate game." The game gives me a feeling for the probability of the particular plate of the car in front of ours, measured against all the possible car plates, a question I feel compelled to answer because, for some mysterious reason, those plates look almost always "peculiar." (To enjoy the game, your students don't have to be mathematical geniuses like the Indian mathematician Ramanujan, who was once told by one of his English mentors, "The license plate of the taxi you took to come here exhibits a 'most indifferent' number." "Not so," exclaimed Ramanujan, after just a glance at the plates, "It is the only integer that can be expressed in *two* different ways as the sum of the squares of *two* integers.")

Ask your students,

PROB

What is the probability that the three numbers on a common type of car plate in the United States be three *different* numbers? What is the probability that the three numbers be *not* all different?

This type of plate can be represented by the symbol:

$$N_1\ N_2\ N_3 - L_1\ L_2\ L_3,$$ ⑪

and you should assume that both numbers and letters are assigned to plates at random and, hence, all have the same probability of being chosen. Assume also that plates starting with zero or having all zeros are acceptable (although in the United States they are not). Do not consider vanity plates (e.g., ILOVEYOU2).

COMB

Assume that you have 3 separate sets of the 10 numbers 0 to 9, and choose the first number for the plate out of the numbers of the first set, the second out of the second set, and the third out of the third set. You then have

$$10 \times 10 \times 10 = 10^3 = 1{,}000$$ COMB

ways of combining the numbers of the 3 sets to get the 3 numbers of your plate. Hence, all the possible combinations of 3 numbers in a plate are 1,000. (In this case, you can simply count 1,000 numbers from 000 to 999, yet the general counting technique is important as follows.) Similarly, if you have 3 sets of the 26 letters of the English alphabet, you can combine them into groups of three in

$$26^3 = 17{,}576$$

ways. Suggest to your students they check that, therefore, one can make 17,576,000 different car plates of the type we are considering.

To build up a plate with 3 different numbers from the same set, you can take any one number from the first set in 10 different ways, but then you are left with only 9 numbers in the second set unequal to the first number and only 8 numbers in the third set unequal to the first two numbers you have chosen. Thus, you have

$$10 \times 9 \times 8 = 720$$

possible combinations of three different numbers, and the probability of one such plate among the 1,000 possible plates is 0.72.

Since the sum of the first and second probabilities must add up to 1, the probability of a plate not having 3 different numbers is only

$$1 - 0.72 = 0.28.$$

Of course, you may check this result by computing first the probability of a plate having 2 equal numbers and then of one having 3 equal numbers. The first can be obtained in $3(10 \times 9 \times 1) = 270$ ways, because the unequal number can be in 3 positions, with a probability of $270/1{,}000 = 0.27$; and the second in $10 \times 1 \times 1 = 10$ ways, with a probability of $10/1{,}000 = 0.01$. Their sum is 0.28, as more easily evaluated above.

COMB

How many different plates can you get with 3 different numbers *and* 3 different letters?

The combinations of 3 different numbers are 10 × 9 × 8 = 720; of 3 different letters, 26 × 25 × 24 = 15,600; and the number of plates is 720 × 15,600 = 11,232,000. Your students may be surprised there are so many. Ask them, "What is the probability of one such plate?" (It is 0.0639 or about 1 in 16.)

How many different plates can you get that have 3 equal numbers and 3 equal letters? What is the probability of one of them?

The combinations of 3 equal numbers are 10 × 1 × 1 = 10; of 3 equal letters, 26 × 1 × 1 = 26. You can put together 10 × 26 = 260 plates with 3 equal numbers and 3 equal letters. Since there are 17,576,000 possible different plates (see the first question in this game), the probability of one plate with 3 equal numbers *and* letters is 260/17,576,000 = 0.0000148, or 1 in 67,600.

How many "symmetrical" plates can you put together of the type

$$N_1\ N_2\ N_1 - L_1\ L_2\ L_1?$$

Once you choose N_1, something you can do in 10 different ways, there are 9 ways for N_2 to be different from N_1 but only 1 to be equal to N_1. Hence, there are 10 × 9 × 1 = 90 ways of obtaining a "symmetrical" number with 3 digits. Similarly there are 26 × 25 × 1 = 650 ways of getting "symmetrical" 3-letter groups. Thus, the number of "symmetrical" plates is 90 × 650 = 58,500, and each such plate has a probability of 58,500/17,576,000 = 0.00333, or 1 in 300.

You can use probability theory with your students for a better purpose than we have so far, so ask them this:

PROB

Is it a good investment to play the New York State Lotto in the hope of hitting the big jackpot and making millions of dollars?

In this betting game, 50 balls, numbered 1 to 50, are put into an urn. Then 6 balls are taken out of it, one at a time. To hit the big jackpot, the 6 extracted numbers must check the 6 numbers you have chosen to bet on.

The probability of the first ball checking one of your 6 numbers is 6/50; of the second checking one of your remaining 5 numbers, 5/49; of the third, 4/48; and so on. Hence, their combined probability is

$$(6/50) \times (5/49) \times (4/48) \times (3/47) \times (2/46) \times (1/45) = 720/11{,}441{,}304{,}000 = 0.0000000629298898097 6,$$

PROB

or 1 in 15,893,198! (Even if you lived to be 100 and played the Lotto 52 weeks a year from birth, your chance would be only 1 in 3,056. I would tell my students that they should find a better investment for their money than playing Lotto in New York State.

I hope that with this kind of problems you will be able to convince your students that probability can be both fun and useful and that it always pays to know what our chances in life may be, even if we will never know for sure whether they will match our probability forecasts.

I would like to end this chapter with the well-known story of the tragic explosion of the space shuttle *Challenger*, which killed seven astronauts in 1986. The explosion occurred because the NASA designers had not checked the loss of strength due to lowering temperatures in the O-rings of the shuttle's fuel tanks. Nobel physicist Richard Feynman proved the influence of low temperatures on the strength of the rings with a very simple experiment: He dropped an O-ring in a glass of iced water, making it so fragile that he could break it with his own hands. No wonder the rings broke up as soon as the shuttle reached the higher layers of the atmosphere!

Mentioning to your students these and similar catastrophes due to forgotten and incomplete checks in engineering design may impress them with the basic importance of checking the answers to their own problems, thus helping them to acquire a "checking habit" to their own advantage and, perhaps later on, to that of all of us.

5

The Logic of Error

ETHAN AKIN

Sometimes you hear something surprising, then you suddenly notice you're not surprised. In the early 1970s a counselor at the City College of New York organized a discussion group for students who were afraid of mathematics. To several of us in the Math Department he described his students' view of our subject: They saw mathematics as a disorganized mess of unreasonable rules. Each rule appeared unrelated to its predecessors, and all were rendered unreliable by mysterious exceptions. In short, they thought of mathematics patterns the way I think of irregular verbs in a foreign language. Shocked at first, our impulse was to cry out: "No, math is not like that. Math is beautiful, neat, and logical." However, we became uneasily aware that their descriptions were merely making explicit attitudes that were already familiar to us. This view of mathematics as muddle was a mistake with which we did daily combat.

It took much longer before I noticed that I myself was making an analogous mistake about the students' patterns of thinking. I began to realize that their errors were not entirely chaotic. Underneath the designs of correct logic, their errors were organized in layers of attractive, wrong rules and natural, if incorrect, patterns of thinking. The perceived chaos resulted, in part, from the discordance between the patterns—the clash between the false rules and the true.

Looking for patterns in the students' thinking made me much more self-conscious about my own. It became apparent that the high-level thinking used in research mathematics, and which we are trying to teach as "problem solving," uses other, almost unnoticed, skills. Especially the symbolism and notation in algebra is packed with subtlety to which familiarity blinds us. For example, I cannot remember when I learned the routine for long division of numbers. I do remember that some teacher told us that this was a very advanced topic in the Middle Ages. Only when I began teaching did I try to imagine a long-division problem in Roman numerals and appreciate the power of the Arabic number system and the associated arithmetic routines.

I don't really understand much about this buried psychology of logic, although I believe it is well worth excavating. What I present below is a collection of fragments, samples of mistaken rules, or examples where symbolism exerts a force supplementary to the mathematics. With these as samples, I recommend that you look out for and collect similar mental tangles, for these can often be the simple blockades to successful problem solving. These often overlooked deterrents need to be addressed before the student can embark on the path to successful problem solving.

Everything Is Linear

I commence with what I think of as the most popular wrong rule, which we have all seen in operation and which appears in the following examples:

ALG

$$\sqrt{(a+b)} = \sqrt{a} + \sqrt{b}$$

TRIG

$$\sin (a+b) = \sin a + \sin b$$

Of course, students don't usually quote such a rule; they just apply it as in

$$9x^2 + 1 = 4,$$

where they take the square root of both sides:

ALG

$$3x + 1 = 2$$

Or there is the dual version:

$$\sqrt{(x+1)} + \sqrt{(2x-5)} = 3,$$

in which squaring eliminates the radicals:

$$(x+1) + (2x-5) = 9.$$

My students tend to recognize and accept this rule as we go over the problem: "You square both sides, right? So we get the square of this plus the square of that equals the square of 3, right?" Right.

This pattern is so seductive that you have to confront it. I tend to talk almost as much about this false law as about the true rules for manipulating radicals. A more advanced, but similar, example comes up in matrix algebra.

The most important rule in dealing with matrices is not the associative or even the distributive law but the commutative law for multiplication: $AB = BA$. It is so important because it is not true and because students use it anyway.

Of course, it is not always clear that the benefit of describing a wrong rule outweighs the risk of implanting the rule without the accompanying "Thou shalt not." However, in my opinion, the wrong rule is sometimes a good introduction to the right one, as demonstrated in the following section.

How Do Fractions Add?

I recall a book in which the author sneers at a student who adds fractions according to the rule

$$\frac{a}{b} + \frac{c}{d} = \frac{a+c}{b+d}.$$

To me, this seems a very sensible rule. I introduce fractions by inviting students to invent the rules that they would like to use. They usually adopt this rule and its analog for multiplication. I can then suggest that multiplication is easy, because the natural rule is the correct one. This raises the question of why we do not use for addition the perfectly sensible rule shown above.

You do not have to think much about the meaning of multiplication of fractions, because it has a rule you could make up in a dream and use in your sleep. Addition of fractions is a nightmare by contrast and requires some serious contemplation.

ARITH

Notice first that there are two ways to read $\frac{2}{3}$: as "2 over 3" or as "2 thirds." It is the second which is the key to fraction addition. In every language I know of, different words are used for cardinal and ordinal numbers: "three and four," as distinct from "third and fourth." This distinction fits together with the familiar pie-cutting illustration of fractions. The numerator of the fraction is the count, the number of pieces of pie. The denominator is what you are counting, the size of each piece. When you think of the denominator as the unit in which you are making a measurement then the strange computation

$$\frac{2}{3} + \frac{5}{3} = \frac{7}{3} \quad \cdot$$

appears perfectly sensible when rewritten as

2 thirds + 5 thirds = 7 thirds,

no more peculiar than

2 meters + 5 meters = 7 meters.

The full routine of fraction addition, if no easier to perform, becomes easier to understand. Just as we convert to a common unit such as centimeters when adding measurements in feet to those in meters, so also do we convert to the common denominator sixths when adding thirds to halves.

(There is one exception to the observation that multiplication and division of fractions are simple. There is a strictly linguistic problem revealed by asking quickly, "What is a third of six?" "What is six divided by three?" "What is six divided by a third?" and getting "Two" each time. In common speech the third phrase is sometimes used [wrongly] to mean the same as the other two.)

Order-of-Operations Nonrules

ARITH

In reteaching for algebra the order of operations rules from arithmetic, it is important to note that the conscious use of the rules is only a temporary expedient. For example, consider the computation problem

$$3 + 2 \times 5.$$

You can arrive at the correct 13 and avoid the error of 25 by consciously applying the rule that multiplication takes precedence over addition. But the instinct that leads to the 25 is really the right one. The way the problem is printed looks strange to anyone who is accustomed to algebra or even arithmetic, because once you have digested the order-of-operations rules, you no longer think about them. Instead, you see them. For instance, you see the expression

$$3 \times 2 + 7$$

as though it were written

$$3 \times 2 \quad + \quad 7.$$

In other words, you see a (+) or (×) sign as forming a linkage between symbols like a chemical bond, and the (×) linkage is much tighter than the (+) link.

The analogy I use is playing a musical instrument. When the playing is correct and smooth, you do not consciously direct your muscles. The knowledge is "in the fingers." Only in correcting an error do you slow up and take conscious control, practicing and repeating to retrain the fingers to get the passage right. As the error is corrected, you gradually relax the conscious control until finally the corrected portion is played fluidly. With order of operations the situation is similar. You don't practice to use the rules; you practice to train the eye so that you no longer need to use them.

An even better example, where the symbolism is intended to invite an unconscious blurring of meaning, occurs in the following section.

The Many Minuses

ARITH

For many years, I taught that the symbol (–) had two different meanings: subtraction and, when attached to a number, the negative sign. I don't remember how long it was before I noticed the third meaning of the symbol. I think I had to discover it on my calculator, where it is designated "chs" for "change sign." The confusion between this unary operation and the sign of a negative number accounts for the common mistake of thinking that $-x$ is negative. (This causes incredible confusion when you attempt to define the absolute value function formally by using a branching formula, or "$-x$ when x is negative.") Especially because a calculator must use different symbols to distinguish them, the questions arise: Why use the same symbol to refer to these three different things? Is this not proof at last that mathematicians are engaged in a sadistic conspiracy to make their subject harder to learn? Actually no. The intention is to allow you to switch meanings without thinking about it.

While it is hidden by the symbolism, subtraction of signed numbers works like division of fractions. Just as you "invert and multiply" to divide, you "change sign and add" to subtract. That is, to perform

$$(-3) - (-5),$$

what you really do is change the sign of minus 5 and replace the subtraction by addition—that is, replace the original problem by this equivalent:

$$(-3) + \text{chs}(-5) = (-3) + (+5).$$

But the whole point is to avoid for subtraction of signed numbers what is unavoidable for fraction division. You should think of the only operation as addition. Each minus is to be regarded as multiplied onto the term that follows it. Thus, the problem above should look (unconsciously) as though it were written

$$(-3) + - (-5).$$

Then use "minus times a minus is a plus" and proceed.

Again the question arises whether it is causes more confusion to point out the three meanings than it does to shift from one use to another without comment. For an algebra class, I favor the former if for no other reason than to avoid the "$-x$ is always negative." mistake mentioned above.

There are also cases where the teacher operates unconsciously in ways which cause some students to lose the thread of the argument.

The Right Rewrite

When you solve for y in the equation

ALG

$$2y = 3x + 4,$$

there are at least three different ways to write the answer:

$$y = \frac{(3x+4)}{2}$$

or

$$y = \frac{3x}{2} + 2$$

or

$$y = \frac{3}{2}x + 2$$

We choose which one to use based upon our subsequent intentions. Thus, if we are about to substitute values of x, we are likely to use the first or second form. On the other hand, if we are computing the slope we use the third form, which displays the coefficient of x. These equivalences are not obvious. For example, second-term calculus students often fail to recognize that

$$\int \frac{dx}{x+1}$$

and

$$\int \frac{1}{x+1}\,dx$$

are the same. I have found it hard to teach this facility for choosing the appropriate form. The best I can do is to point out the various alternatives as they happen and indicate the advantages of each form. However, this leads to the classic question posed in the next section.

But Which Is the Right Way?

This is a problem that I think we in the teaching profession have imposed on the students. Many of them have picked up an Aristotelian view that there is an absolute right form that expressions should be transformed into, whereas, of course, the rightness of a particular form depends on the use you are going to make of it. Should fractions be combined into one large one or broken up into separate, smaller pieces? Should fractions be reduced to lowest terms or built up to a common denominator? Sometimes one form is needed and sometimes another.

One way we give this impression is by using the instruction "simplify." What we mean is, "Part of this problem is to recognize from its form what is wanted based upon our recent work. To be more explicit in my instructions is to provide a hint, which should be unnecessary. After you have decided what to do, do it." However, the student interprets the instruction as "Write this expression in its best form."

For example, should a polynomial be multiplied out or written in factored form? It depends. The best example of this comes up with curve sketching in first-year calculus. Because differentiating a sum is easier than differentiating a product, you often multiply out to take the derivative. But the next step in curve sketching is to look for zeroes of the derivative in order to see where it is positive and where negative; for this you factor. Then you repeat the process for the second derivative. Thus, for a number of such problems, you go back and forth, first factoring and then multiplying out, depending at each step on what you need to do next. You want to do the right thing, but which thing is right depends on what you are doing.

The Asymmetry of Equals

The distributive law, $a(b+c)=ab+ac$, is a beautiful example where the psychology and logic of the symbolism separate. The equation is symmetric. The left side equals the right. But as transformation rules, moving right and moving left feel so different that they are given different names. It even requires a moment's pause to see that the two processes are just reverses of one another. The students are first given a page of problems where the question is the left side and the answer is the right (multiplying out or clearing parentheses). Then the next section consists of the reverse (factoring). It is useful to point out that one process is not merely a "check" for the other but that the two are really just one equation viewed two different ways.

ALG

An analogous problem arises for the fraction equation

$$\frac{a}{b} \bullet \frac{k}{k} = \frac{a}{b}.$$

Moving left to right is "reducing fractions" by canceling a common factor from the numerator and denominator. Moving right to left is "building fractions," the central step in fraction addition.

It is precisely the fact that these two processes seem to many students unrelated that accounts for a characteristic error in fraction addition that would be funny if only it did not happen so often. The student recognizes that a common denominator is necessary, computes the least common denominator correctly, and furthermore converts each summand to the common denominator. Then, on the brink of success, having completed all of the hard work, the student suddenly notices he or she can reduce the fractions. This the student proceeds to do without realizing that he or she is exactly undoing the work of the previous conversion step. Having lost the common denominator, the student now wanders further into the darkness, adding the numerators and adding or multiplying the denominators.

To See or to Do?
Monomial Versus Polynomial Factoring

Factoring is factoring, after all. But monomial and polynomial factoring not only feel completely different, but the skills developed for the first work against success in the second. For a monomial like

$$30x^3y^2,$$

the student should see the number 30 as containing its prime factors. Furthermore, x^3 is three xs multiplied together, with y^2 similar. Students learn this skill easily, and it is just what is needed for problems like reducing monomial fractions. However, they are then confronted with

$$6x^2 - y$$

and asked, "What are the factors?" They reply, quite reasonably, "2, 3, x, x, and y."

You can point out the difference: In a monomial the only operation is multiplication, while a polynomial contains a mixture of multiplications on one hand with additions and subtractions on the other. It is this mixed state that causes the difficulties in the polynomial case. So the mixed state should be used as a danger signal. While for monomials the factors are just what they appear to be, polynomials are to be factored only by one of the short list of procedures—for example, dividing out the common monomial, the difference of two squares method, or trinomial factoring—leaving other, subtler polynomial factoring skills to be developed at a later time.

However you justify the difference, I fear that the students perceive this as one of the irregular verbs I mentioned in the introduction.

The Abstract, the Concrete, and the Laws

ALG

The original meaning of exponents is wonderfully concrete. Think of x^7 as a bag containing 7 xs all multiplied together. The Laws of Exponents are then the result of observation. For example, the equation

$$x^7x^5 = x^{7+5} = x^{12}$$

just says that with all the xs multiplied, you only have to keep track of the count. So if a bag with 7 xs is poured together with a bag containing 5 xs the result is 12 xs all multiplied together. Similarly,

$$(x^5)^3 = x^{5\times 3} = x^{15}$$

says that 3 bags, each containing 5 xs, contain all together 15 xs. Trouble arises when you want to define negative or fractional exponents. How should you multiply together minus 3 xs to get x^{-3}? On the other hand, multiplying half an x seems perfectly simple and so $x^{½} = x/2$.

It is worth pausing here to talk a bit abstractly about the process of abstraction. The problems the students have, illustrated by the above questions, arise from trying to extend the original definition of exponent by extending the original physical sense of what exponents are. However, we are at a fork in the road, and we should pause before we decide which way to go. From our original physical picture of exponents we observed some useful rules, the Laws of Exponents. In fact, they are so useful that when we extend our notion of exponents to negative and rational numbers, we want to make sure that we retain the rules even at the cost of our original physical intuition. That is where the peculiar new definitions come from. Think of x^{-3} as an unknown quantity. If it satisfies the Law of Exponents, then we must have

$$x^{-3}x^4 = x^{-3+4} = x,$$

and so we see that

$$x^{-3} = x/x^4 = 1/x^3.$$

Similarly,

$$(x^{½})^2 = x^{½\times 2} = x$$

implies that $x^{½}$ has to be the square root of x.

There is a similar change in meaning when we abstract the definition of "vector." Originally, a vector is defined as an object having a magnitude and a direction, in contrast with a scalar, which is just a magnitude. Vectors are pictured as little arrows. The physical motivations for the idea are force and velocity. We discover notions of addition and scalar multiplication for vectors that relate naturally to phenomena concerning forces and velocities. We then observe that these operations satisfy a list of useful rules analogous to those of ordinary algebra. In linear algebra, these ideas are abstracted to define a "vector space." Again, the physical motivation is thrown away. What is retained is the list of algebraic rules as the axioms for a vector space.

Zero Versus Nothing

ALG

This is probably my favorite example, because it illustrates so well the gap labeled "mathematical sophistication." Even very early algebra students find the equation:

$$5x = 3x + 6$$

easy to solve. However, even calculus students are often unable to correctly solve

$$5x = 3x$$

I am still not sure why. Sometimes I think it is symmetry that prevents an "obvious" choice of side on which to isolate x. It is, for example, true that the equation

$$5x + 8 = 3x + 6$$

is harder than the first equation. But it is not nearly as hard as the second. Another possibility: In the first you have the 6 from which to take away the $3x$. You have something against which to brace your foot, as it were. Whereas in the second there is nothing there.

The 10 difficulties described above illustrate what seems to me to be an important and wide-spread problem. Mathematical symbolism, even at the very elementary level, has a subtlety and power that we use without conscious thought. Like an ideal servant, it performs its tasks unobtrusively. As working mathematicians and problem solvers, we are appropriately unconscious of its work. However, as teachers of mathematics and problem solving, we can help our students if we pay attention to its hidden power.

6

Trial and Success

FRED PAUL

Today's students—especially those in mathematics classes—seem to want and expect immediate and correct solutions to all their questions. Unfortunately, this attitude does not help in the development of problem-solving techniques that involve constant trials and questioning. The trial part includes the willingness to go down blind alleys, try alternatives, and accept disappointments—in short, sometimes to be wrong. The expression "trial and error" carries an especially negative connotation for the many who often won't even try, for fear of being wrong. It is essential that the so-called wrong paths in problem solving are treated not as errors but as often necessary steps toward a solution. They can give important clues and even provide insight, sometimes in totally unexpected directions that lead to an answer to a different question which can be useful at another time. Why not, then, consider the process as one of "trial and success"?

A few simple strategies depend on the instructor's ability to guide the problem-solving process as a group endeavor, promoting constant questioning and interaction with as little interference as possible. One of the strategies used to encourage this kind of reaction is to give no response of any kind to an answer, whether it is correct or not. In other words, when students offer responses to a question in the course of discussion, occasionally the teacher would give no indication, verbal or facial, as to the correctness of the statement. What ensues is often quite interesting. A student will very often change the response, implying either not being sure of what was said or feeling confusion about not getting immediate confirmation. In either case, this will usually elicit contributions from others in the group either to support or to challenge the original response. At least two purposes will have been served. First, discussion has been initiated. But a second, possibly more important aspect must be considered: When a teacher acknowledges a correct response from a particular student, the thought process of that student concerning the item in question may end—that student got it! However, that correct response being confirmed may stop the thought processes of many of the other students who have not arrived at that point of understanding. These other students could become lost in the continuing problem-solving process.

GEOM

⑦

As a specific example, ask a group (composed of students or teachers) how many surfaces there are on a cube. The correct answer will come rather quickly and confidently, but the reactions usually become hesitant when the instructor does not immediately verify the answer. It can and does become quiet (fear of error) until some indignant insistence or even some different answers are offered. After group members reach agreement on the answer, ask them how many surfaces there are on half a cube, and make a vertical slicing motion. At this point, the instructor must become a true discussion leader, encouraging the opening up of divergent thinking. Of course, many answers are possible. Six is the most obvious answer, but five is also fine (a diagonal cut). It shouldn't be surprising to hear a quick answer of three before a certain realization occurs, but it might be surprising to hear answers greater than six. These are possible when different ways of cutting the cube in half (for instance, in stairs) are considered; in this way, any number of surfaces becomes possible, as well as different kinds of surfaces. This can lead to a truly mathematical discussion of what is meant by "half."

An example like this can help persuade students to try different answers and not worry about being wrong and to use different answers to discuss and analyze the problem. They also keep the instructor alert, since unexpected alternative and correct answers often crop up. Care must be taken not to gloss over such answers. We often predetermine what answer to a question is desired and are not always tolerant of alternatives that may arise from a misunderstanding or different interpretation of a given problem. It is possible that an original or innovative thinker could lose interest if not allowed to pursue avenues that are different or unexpected but possibly legitimate. Contrary to its reputation, mathematics is not always predictable.

Discussions of this kind address the idea of "real-world" problems. It is how problems are perceived, interpreted, and pursued—and not necessarily the topic of the problem or its relevance to a real-life situation—that makes them real world in nature. Most problems from the adult real world have little meaning or interest for students. What have been called "whimsical" problems, while certainly having little to do with the real world, can be useful and practical in the development of questioning techniques and problem-solving skills.

Another strategy to encourage questioning and the willingness to explore and experiment is the use of problems that are purposely ambiguous. Traditional mathematics is notorious for its avoidance of anything ambiguous, but greater and greater emphasis on problem solving in the mathematics curriculum dictates an awareness of ambiguities in many situations. Contrived problems help develop strategies that can be used to solve problems that arise naturally, and lend themselves to a trial-and-success approach. An example is this variation of a familiar classic problem: ⑦

> A fly and a jogger are 12 km apart. The jogger runs toward the fly at a rate of 4 km/hr. The fly travels toward the jogger at the rate of 6 km/hr. When the fly meets the jogger, it turns around and flies back to its starting point, then heads again toward the jogger, then flies back again. The process continues. What is the total distance the fly travels?

GEN

With certain assumptions, this is a very simple problem but one that stumps many. Often, when there is no intimidation and there is a willingness to make guesses, a variety of answers is offered. Eventually (with proper guidance, if necessary), but not always quickly, the ambiguous aspect will be discovered: There is no mention in the problem, as stated, of when the process stops. Most will assume that it stops when the jogger arrives at the fly's starting point, but this must be stated in the problem in order to avoid the ambiguity. Even with this done, the problem is not always solved quickly and can be the source of a good exchange among students. It is not just in geometric problems that students must be reminded to consider all the information given. Because of the constant activity of the fly, going back and forth, and because it is the fly's distance traveled that is being sought, students usually concentrate all their attention on it. Actually, by ignoring the fly and working with the jogger, the time of travel is easily arrived at. By using nothing more than the formula rate × time = distance, it can be determined that the fly traveled 18 km.

Many mathematicians will object to the purposeful use of ambiguity in a problem, but, while best avoided in a testing situation, it is a useful tactic to encourage a diversity of questions and to make students aware that problems do not as a rule appear in neatly wrapped and predictable packages. It might be added that many of the same people who would object to the ambiguities would consider mathematics in many ways to be a guessing game.

Another important aspect of a problem like the jogger-fly example is that it demonstrates that difficulty is not a proper measure for a useful and valid problem. Whether easy or difficult, routine or nonroutine, the good problems are those that will nurture questions, diverse thoughts, differences of opinions, guesses, and, above all, a willingness to pursue various alternatives and keep going if unsuccessful. This is the true trial-and-success pattern, the joy of the problem-solving process. A solution is merely one other part of this process and could and should be a catalyst for further questions, discussion, and analysis. In this way, mathematics becomes the vital and interesting subject that truly helps students learn to think.

We have in the following example a routine problem with some nonroutine features:

GEOM

Three circles are mutually tangent, and each is tangent to the same line. Two of the circles are congruent and the third, incongruent circle has a radius of 3. What is the radius of the congruent circles?

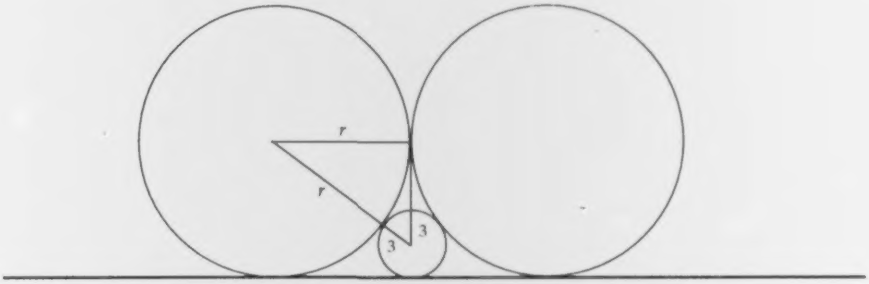


Figure 6.1

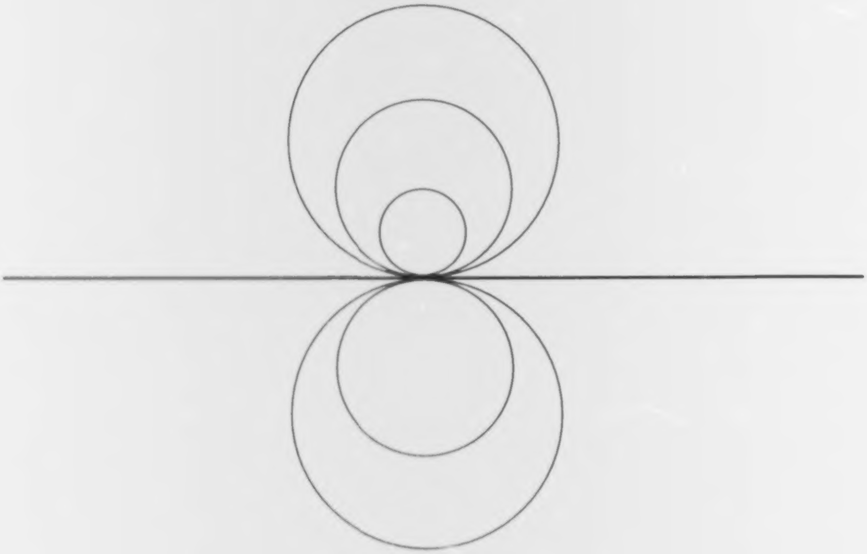


Figure 6.2

As with most interesting geometry problems, the solution depends on the proper diagramming of the given information. This can be the source of a valuable full-class discussion of which most teachers don't take advantage. Again, it is essential for the students to be encouraged to consider different possibilities with a constant exchange among them. The feature of this problem, with respect to the trial-and-success mode, again involves an ambiguity. There is a unique answer when the three circles are on the same side of the tangent line (Figure 6.1). The answer involves the use of geometric and algebraic properties. As shown in Figure 6.1, the drawing of three lines—one from the point of tangency of the two congruent circles to the center of the third circle, and two from the center of either congruent circle to the center of the third circle and to the point of tangency of the two congruent circles—will form a right triangle with sides r (radius of the congruent circle), $r - 3$, and $r + 3$. Using the Pythagorean theorem will give a radius of 12 for the answer. There is no limit to the answers if the congruent circles are on opposite sides of the tangent line (Figure 6.2). Problems like this—basic and easy enough for all students to understand and discuss, yet challenging and not trivial—promote the searching and questioning attitude needed to get at the essence of mathematical problem solving for all students at all levels in preparation for the everyday problem solving they will face as adults.

An example that students can have fun with, and that also emphasizes the trial-and-success approach, is the following problem:

> Jim asks a classmate how old she is. The classmate answers, "I was 14 the day before yesterday, but I'll be 17 next year." What are the dates of when the classmate is talking and of her birthday?

GEN

Student reactions to questions of this type can range from puzzled concern to outright denial of the possibility. However, questions and guesses will point out the importance of the students' ability to communicate and work together and demonstrate the idea that this is what problem solving is all about. There is nothing ambiguous about this problem, with a December 31 birthday being talked about on January 1. The problem solver may make many guesses or "errors," however, before arriving at the correct response or "success."

Good problems can also be used to expand and illustrate major mathematical concepts:

> A truck carrying 4,000 crates of merchandise cross-country drops off half the crates at its first stop. Only full crates may be dropped off. It drops off half the remaining crates at the second stop. Half the remainder is dropped off at the third stop. If the pattern continues, at which stop will the last crate be dropped off?

GEN

The ensuing discussion will most likely involve fractions (part of the crate, even though the problem avoids this) and possibly the need for more information. Ultimately, however, there should be agreement that, no matter what, half of what ⑦ is in the truck must remain there after each drop-off point. There is no "last" crate to drop off. The "trial" in this situation will be the various answers offered, and the "success" will be the realization that there is no answer. Problems like this are valuable since they lead to a different kind of answer and allow the instructor and the class to delve into other topics such as limits and infinity.

The questioning and discussion resulting from problems like these often should be considered in a full-class situation. Many teachers will claim that there isn't enough time for much of this kind of activity because of curricular and testing requirements—but actually there is not enough time *not* to do it. The ability to think through and understand the problem-solving process is arguably the primary goal of mathematics instruction. It is necessary for students to hear their questions, answers, and comments clarified, interpreted, and challenged by other students. This exchange of ideas will become valuable for later tasks. It is equally important for the teacher to monitor the class discussion. Many who are not necessarily "good" mathematics students in the traditional sense—that is, those who don't score well on tests and may have little interest in the mechanical or rote aspects of the subject—may show original thinking and insight in the nonroutine problem-solving situation when given the opportunity in an open-ended discussion. In this way, a teacher can discover abilities not otherwise apparent, and the student can develop a more positive attitude toward mathematics. These students may then view the subject as more than a set of rigid rules where the idea of any kind of error denotes failure. This prospect further emphasizes trial and success.

A full-class approach to problem solving also allows for examples containing topics and terms not familiar to everyone. For example, there are those who will frown on the use of a problem involving sports, since there will be some in the class who will not be familiar with the given sport and its terminology. In adult life, we often face problems that include unfamiliar aspects; an important part of the problem-solving process is the clarification of language and definition of terms. The full-class situation is ideal for pursuing these matters and stressing the importance of a clear understanding of what is given and what is sought. An example:

GEN

In a recent tennis tournament there were 61 players entered. In the first round and in any later round involving an odd number of players, one of the players had a bye. Every match continued until one player won with the loser eliminated. How many matches in all took place until a single player was left unbeaten?

Before a solution can be attempted, all terms have to be understood by all students. If asked to work individually, some would have no idea how to proceed. A quick review of tennis might be needed; most likely "bye" (a player being excused from having to compete in a given round) will need to be explained. It is hoped that by doing this, students will become willing to try problems regardless of their familiarity with the topics. Once that is out of the way the discussion centers on ways to solve the problem, and the trial and success is in operation. In ⑥ this particular case, many may actually make a diagrammed schedule of the matches and count up the games played, and that is fine. Of course, no matter how ⑬ many players are entered, all but one must lose, so the number of marches played is one less than the number of players entered—in this case, 60.

This type of activity must become a regular part of almost every mathematics course, even if it means extending the usual time allocated to complete specific goals and outcomes. For the most part, this extension could be avoided with a shifting of certain emphases and the use of technology to streamline the traditional offerings. The many schools that now extend courses by slowing down the presentation of material would do well to pay more attention to the development of problem solving that might provide more meaning and motivation for the students.

One of the important features of the extensive use of so-called nonroutine problems with a trial-and-success approach and in a full-class context is the development of students' awareness of the wide application and usefulness of mathematics. Not everything is learned in order to succeed on a test question. In fact, most questions such as the ones considered are not usually found on a classroom test because they require discussion and trials, which usually entail more time than can be allowed for most examinations.

Furthermore, part of the problem-solving approach involves having time away from the problem in order to consider and reconsider different possibilities. Many problems become clearer when a student can "take it home and sleep on it." The entire learning experience is enhanced by students realizing and accepting the fact that all problems will not be solved by everyone, that some will be solved by none, that some will have several routes to common or alternative solutions, and that some will have no solutions at all. Students will soon realize that most problems are not solved in isolation. Talking to different people (fellow students, parents, siblings, friends, etc.) and using other resources become valuable parts of the process but not part of the usual examination process. With this understanding and concept development, the common exercises and problems on classroom tests often become easier and sometimes quite trivial.

The problems discussed and the wealth of additional possibilities can be approached individually or in small groups, exemplified by the resurgence of what is today called "cooperative learning." Although both of these approaches are valid and useful and should be important parts of the teaching of problem solving, a full-class approach must not be ignored. At times, the technique will involve preassigned problems, the discussion of which will be about work and thinking that students have already done. It is also important that problems be tackled by full class participation from the beginning. In this way, all students get the benefit of being part of the entire process, contributing to and hearing different ways of interpreting a problem and determining how to reach a solution. Discussions of this nature become valuable tools for the instructor's assessment not only of student performance but also of the mathematics program itself. Even for the teachers and for the instructional validity of the program, it is a matter of trial and success.

7

Reduce, Expand, and Look for a Pattern

STEPHEN KRULIK

JESSE A. RUDNICK

Rarely in problem solving does the problem solver resort to a single strategy. In most cases, two or more strategies are combined to resolve the situation. In fact, reduction and expansion almost always depends on discovering a pattern. Ironically, many years ago, the mathematician W. W. Sawyer described mathematics as a search for patterns.

In order to fully appreciate the strategy of reduction and expansion, we must first differentiate between simple reduction and reduction and expansion. In many textbooks, the strategy of reduction is referred to with the direction "solve a simpler problem." When using this strategy, the numbers in the problem are reduced to simpler quantities, the simpler problem is solved, and the solution procedure is then applied to the original version of the problem. On the other hand, reduction and expansion is an appropriate strategy to use when the problem contains arbitrarily large numbers and the solution is virtually independent of these numbers. When using this strategy, one reduces the arbitrary number to the simplest possible case (usually 0 or 1), notes the results in a table or chart, and then permits the number to expand. Again, the results are noted in the table, until an underlying pattern emerges for which a generalization can be made.

Teachers should not treat these tables and charts in a casual manner. Children need experience in handling data. The reduction-and-expansion strategy almost always involves the use of a table to keep track of the changes that occur as the variable expands. These tables should be designed and completed by the children. A careful analysis of this table permits the student to observe the developing pattern. This analysis will serve students well in all of their mathematical experiences.

To illustrate this strategy, we will present a series of problems. Note that problems are merely vehicles for the teaching of reasoning and problem solving. They need not be from the real world nor applicable to any particular occupation. It has been our experience that children can easily be motivated by science fiction, fantasy, or recreational topics.

Example 1

> Laura is training her pet white rabbit, Ghost, to climb a flight of 10 steps. Ghost can hop up 1 or 2 steps each time he hops. He never hops down, only up. How many different ways can Ghost hop up the flight of 10 steps?

In this problem, the numbers 1, 2, and 10 are arbitrary. However, it is the magnitude of the 10 that suggests using reduction and expansion—after all, if the number of steps had been smaller (say, 3 or 4), we could simply have written out all the possible ways. Thus, we reduce the number of steps to 1 and proceed to note the results as the number of steps expands to 2, 3, 4, and so on, through the counting numbers. This expansion needs to continue only until a pattern has been established. Care must be taken to ensure that the pattern is the one underlying the problem; do not generalize too quickly!

We make a table to record our data (see Table 7.1).

After the fifth step, it becomes apparent that the Fibonacci sequence[1] was the number pattern underlying the problem. Care must be taken to ensure sufficient trials to establish the correct pattern. Note that some people who stopped after steps 1, 2, and 3 would have selected the counting numbers as the pattern and would have gotten 10 as their answer. For those students who are still wary, the procedure has been established to continue in this way, enabling them to verify the pattern.

We have found the answer to the problem based on our conjecture that the Fibonacci sequence is the underlying pattern of the problem. Depending on the mathematical sophistication of the students, an attempt should be made to justify

Table 7.1

Number of Steps	*Number of Ways*	*Ways*
1	1	1
2	2	1-1, 2
3	3	1-1-1, 2-1, 1-2
4	5	1-1-1-1 2-2 2-1-1 1-2-1 1-1-2
5	8	1-1-1-1-1 2-1-1-1 1-2-1-1 1-1-2-1 1-1-1-2 2-2-1 2-1-2 1-2-2
6	13	
7	21	
8	34	
9	55	
10	89	

the conjecture. The fact that the rabbit takes only 1- or 2-step hops implies that Ghost can only reach the 5th step from either the 3rd or 4th step. Thus the total of these possibilities (step 3 and step 4) is the entry in the box for the 5th step. This exhibits the notion seen in the Fibonacci sequence, where each term after the first two is the sum of the previous two terms.

The main purpose for this problem was to involve the students in analysis, reasoning, and the reduction-and-expansion strategy, but more should be done. The problem should be used to stimulate further discussion. Creative teachers would ask "what if" questions: What if the rabbit could hop up 1, 2, or 3 steps at a time? How would this affect the answer? Would the number of ways increase or decrease? Using the same reduction-and-expansion technique reveals interesting patterns, including a Fibonacci-type sequence with three consecutive terms being summed.

A top priority for all mathematics teachers at every level is to reveal the power of mathematics to their students—that is, to use mathematics to completely describe a situation or a phenomenon and to predict with certainty the outcome. This goal can often be achieved by generalizing after reduction and expansion have revealed the fundamental pattern.

Table 7.2

Tie Score	*Number of Scores*	*Scores*
0–0	1	0–0
1–1	4	0–0 0–1 1–0 1–1
2–2	9	0–0 1–0 2–0 0–1 1–1 2–1 0–2 1–2 2–2
3–3	16	0–0 1–0 2–0 3–0 0–1 1–1 2–1 3–1 0–2 1–2 2–2 3–2 0–3 1–3 2–3 3–3
⋮	⋮	
8–8	81	

Example 2

At the end of the 7th inning of last night's baseball game, the score was 8-8. How many scores were possible at the end of the 6th inning?

Once again the arbitrary numbers 8-8 complicate the problem. Thus we reduce the score to 0-0, expand it to 1-1, 2-2, 3-3, and so on, keeping track of the results as we go. Again we make a table (see Table 7.2).

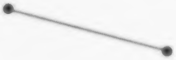
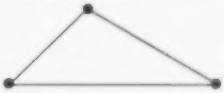
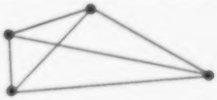
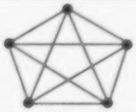
We can generalize the set of perfect squares revealed in the table to $(n + 1)(n + 1)$, or $(n + 1)^2$, where the tie score was n-n.

This reveals the unique power of mathematics to describe a situation. Regardless of the size of the score, a tie produces the expression $(n + 1)^2$.

Again, the creative teacher can ask, What if the score was a-b, not a tie? This leads to the expression $(a + 1)(b + 1)$. Thus the tie score n-n is now a special case of the general expression, where $a = b = n$.

Some students may recognize that this is an illustration of the counting principle. Since the first team can have any of nine scores (0-8) and the second team can also have any of nine scores (0-8), there will be 9 × 9 or 81 possible scores.

Table 7.3

Number of Offices	*Number of Lines*	*Drawing*
1	0	
2	1	
3	3	
4	6	
5	10	
. . . 26	. . . 325	

Example 3

> There are 26 teams participating in the annual football draft. Each team office has a direct line to each of the other team offices. How many telephone lines are there?

This problem is a variation of the well-known "handshake problem." Some students may be tempted to either make a physical model (using string and pegs) or draw the diagram. The 26 offices make these approaches impractical, although either of them could be used to find the answer. However, the problem is a perfect example for using the reduction and expansion strategy: Reduce the number of offices to 1 and then expand until a pattern is observed. Include in the table a drawing of the action.

We have placed the column with the number of telephone lines directly adjacent to the column with the number of offices. Although this is different from the order suggested by the problem, it has been our experience that close proximity of the numbers makes discovery of the relationship easier. Notice that a problem that begins with telephone lines results in a geometry problem involving the sides and diagonals of a polygon of *n* sides.

Some students may require carrying the table further, in order to observe the pattern (i.e., the difference between the terms in the "Number of Lines" column is the counting numbers 1, 2, 3, . . . 25). Many students will continue this for all 26 terms. However, some more astute students may observe that the number of telephone lines required for *n* offices is given by the expression

$$\frac{n(n-1)}{2}.$$

This is not a simple observation to make. The teacher may have to lead the student to this discovery.

For senior high school students, this is an excellent opportunity to consider the method of finite differences. This can be demonstrated as shown in Table 7.4.

Table 7.4

Number of Offices	*Number of Lines*	Δ_1	Δ_2
1	0		
		1	
2	1		1
		2	
3	3		1
		3	
4	6		1
		4	
5	10		1
		5	
6	15		
.	.	.	.
.	.	.	.
.	.	.	.
	.		
	.		

Table 7.5

Size of Squares

Size of Board	*1 × 1*	*2 × 2*	*3 × 3*	*4 × 4*	*5 × 5*	*6 × 6*	*7 × 7*	*8 × 8*	*Total*
1 × 1	1								1
2 × 2	4	1							5
3 × 3	9	4	1						14
4 × 4	16	9	4	1					30
⋮									⋮
8 × 8	64	49	36	25	16	9	4	1	204

Since the constant appears in the Δ_2 column, we know that the formula will be of second degree.

Example 4

How many squares are there on an 8 × 8 checkerboard?

The first reaction to this problem is that there are 64 squares. However, these are only the 1 × 1 squares, and there are, indeed, 64 of them. But what about the board itself? This is an 8 × 8 square. Further analysis reveals that there are 2 × 2 squares, 3 × 3 squares, and so on. Reduction and expansion is the obvious strategy.

There are 204 squares on an 8 × 8 checkerboard. Once again, the more advanced students should continue this problem and develop a formula for the sum of the squares. Table 7.6 illustrates the method of finite differences, revealing that the constant occurs in Δ_3, indicating a cubic relationship. The resulting formula is

$$\sum_{K=1}^{n} K = \frac{n(n+1)(2n+1)}{6}.$$

Table 7.6

	Total	Δ_1	Δ_2	Δ_3
1 × 1	1			
		4		
2 × 2	5		5	
		9		2
3 × 3	14		7	
		16		2
4 × 4	30		9	
		25		2
5 × 5	55		11	
		36		
6 × 6	91			

Another interesting extension might be to ask the students to determine how many of these 204 squares contain an equal number of red and black unit squares.

Example 5

> The new school building has exactly 1,000 lockers and 1,000 students. On the first day of school, the students meet in the schoolyard and agree on the following plan. Student 1 will run through the building and open all of the lockers. Student 2 will then enter the building and close all the even-numbered lockers (2, 4, 6, . . . 1,000). Student 3 will then enter the building and reverse every third locker beginning with number 3 (3, 6, 9, . . . 999)—that is, if the locker is open, he will close it, and if it is closed he will open it. Student 4 will then reverse every fourth locker (4, 8, 12, 16 . . . 1,000). The students will continue in this way, until student 1,000 has reversed locker number 1,000. Which lockers now remain open?

Although this problem could be solved experimentally by using 1,000 coins in lieu of the lockers and using heads and tails to represent open and closed lockers,

Table 7.7

	Locker #	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*
Student	1	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O	O
	2		C	O	C	O	C	O	C	O	C	O	C	O	C	O	C	O	C	O	C
	3			C	C	O	O	O	C	C	C	O	O	O	C	C	C	O	O	O	C
	4				O	O	O	O	O	C	C	O	C	O	C	C	O	O	O	O	O
	5					C	O	O	O	C	O	O	C	O	C	O	O	O	O	O	C
	6						C	O	O	C	O	O	O	O	C	O	O	O	C	O	C
	7							C	O	C	O	O	O	O	O	O	O	O	C	O	C
	8								C	C	O	O	O	O	O	O	C	O	C	O	C
	9									O	O	O	O	O	O	O	C	O	O	O	C
	10										C	O	O	O	O	O	C	O	O	O	O
	11											C	O	O	O	O	C	O	O	O	O
	12												C	O	O	O	C	O	O	O	O
	13													C	O	O	C	O	O	O	O
	14														C	O	C	O	O	O	O
	15															C	C	O	O	O	O
	16																O	O	O	O	O
	17																	C	O	O	O
	18																		C	O	O
	19																			C	O
	20																				C

this is still impractical (what if there had been 10,000 lockers?). Reduce, expand, and look for a pattern!

Observation shows that the lockers whose numbers are the perfect squares remain open. But an interesting question remains. Why the perfect squares? Here is an opportunity to look at some number theory concepts. This problem is based on numbers and their factors. A locker will only be touched by a student whose number is a factor of a locker's number. But factors always occur in pairs (open-close):

12 contains 1, 2, 3, 4, 6, 12

10 contains 1, 2, 5, 10

so all of the lockers should be closed. But why did the perfect squares remain open? Let's look at 16:

16 contains 1, 2, 4, 8, 16

The factor 4 has *itself* as a "mate." Thus, only the perfect squares have an odd number of factors, a fundamental theorem in number theory.

Reduction and expansion is not only an important strategy for problem solving, but it also opens the door to creative mathematical thought. It helps provide students with the mathematical power we are all seeking.

Note

1. In 1202, the mathematician Leonardo of Pisa (known as Fibonacci) published his *Liber Abaci*. In this book, he discusses the now-famous "rabbit problem":

> How many pairs of rabbits will be produced in a year, beginning with a single pair, if in every month each pair gives birth to a new pair which becomes productive from the second month on?

The solution to this problem gives rise to the "Fibonacci sequence," 1, 1, 2, 3, 5, 8, 13, 21, . . . where each term after the first two is the sum of the two terms immediately preceding it.

8

The Pigeonhole Principle for Problem Solving

ALFRED S. POSAMENTIER

WEI LEE

The Pigeonhole Principle

Logical counting has fascinated people through the ages. When the counting is straightforward, we use schemes such as those described in the Introduction. However, sometimes problems ask for an ordering—that is, for the data to be arranged according to certain rules. The inspection of this ordering can best be captured by the Pigeonhole Principle.

Before considering the Pigeonhole Principle in a formal sense, let's consider a few examples that will demonstrate how the Pigeonhole Principle manifests itself.

There are 5 pigeonholes in the tree. There are 6 pigeons flying into these 5 holes. Therefore, there is at least one hole containing 2 pigeons.

If we put 16 flowers into 3 jars, there is at least one jar containing 6 flowers.

There are 2 mailboxes in front of a post office. We want to put 5 letters into these 2 mailboxes. Whichever way we deposit these letters, one of the mailboxes will have at least 3 letters.

It is easy to come up with other examples similar to these three. Consider "letters," "flowers," and "pigeons" as *objects*, and consider "mailboxes," "jars," and "pigeonholes" as *holes*. After careful analysis of the above examples, we should see a simple and important principle: If we have a certain number of objects and a certain number of holes that is less than the number of objects, and we place the objects in the holes, then there must be one hole containing more than one object. ⑱

COMB

To further describe the Pigeonhole Principle, we will consider it in three parts, or subprinciples:

Principle 1. Put $k + 1$ objects into k holes; at least one hole will contain 2 or more objects.

Principle 2. Put $m \geq 1$ objects into n holes. If n divides m evenly—that is, if $m = nq$ $(q \in N)$—then there is at least one hole containing at least q objects.

It follows that if each of the n holes has less than q objects, then the number of objects all the holes contain will be less than nq.

Principle 3. Put $m \geq 1$ objects into n holes. If n cannot divide m evenly—that is, if $m = nq + r$ $(0 < r < n)$—then there is at least one hole containing at least $q + 1$ objects.

It follows that if every hole has fewer than q objects, then the number of objects all the holes contain is less than nq, which is contradicted by the assumption that the total number of objects is $nq + r > nq$.

Principle 1 is a special case of principle 3. Actually, if $m = n + 1$, then principle 3 becomes principle 1.

The Pigeonhole Principle, comprising the above subprinciples, provides a rather effective problem-solving tool for dealing with a specific type of (sometimes) difficult problem.

Examples

On selects 5 cards from a set of cards colored red or blue. Prove that, whichever way you select these colored cards, there are at least 3 cards of the same color.

Proof. Consider the red and blue colors as two different holes. Assume that the red cards are put into the "red hole." Similarly, the blue cards are put into the "blue hole." According to the Pigeonhole Principle, there are at least 3 cards in one hole. This proves that there are at least 3 cards with same color.

This example is simple. But considering colors as holes is usually not easy for students to comprehend at the start. Additional examples are provided to help secure an understanding of this very important and useful principle.

There are n people in a meeting. Every one of them knows at least one of the other $n - 1$ people. Thus, among these n people there are at least 2 people who know the same number of people.

Proof. Since everyone knows at least one of the other $n - 1$ people, then the number of people every person knows is $1, 2, 3, \ldots$, or $n - 1$. Consider "the number of people every person knows" as holes. In other words, we make $n - 1$ holes: ⑯

"knows 1 person" is the first hole;
"knows 2 people" is the second hole;
. . .
"knows $n - 1$ people" is the $(n-1)^{\text{th}}$ hole.

Now we can put these n people, according to the number of people they know, into these $n - 1$ holes. According to the Pigeonhole Principle, there is at least one hole that contains 2 people. This proves that within these n people there are at least 2 people who know the same number of people.

The Pigeonhole Principle can also be used to solve problems related to integer points (or lattice points) on a coordinate system. An integer point is a point with integers as coordinates. If $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ are two points in the X-Y plane, then the coordinates of the midpoint M of the line connecting these 2 points is GEOM

$$x = \frac{x_1 + x_2}{2}, \quad y = \frac{y_1 + y_2}{2}.$$

Clearly, if $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ are integer points, and both x_1 and x_2, and y_1 and y_2 have the same parity, then the midpoint is also an integer point.

Similarly, if $P_1(x_1, x_1, z_1)$ and $P_2(x_2, y_2, z_2)$ are integer points in the three-dimensional plane, and x_1 and x_2, y_1 and y_2, and z_1 and z_2 have the same parity, then the midpoint of the line connecting these points is also an integer point.

> In *X-Y* plane, pick up any 5 integer points. Prove that there must exist 2 of these integer points where the midpoint of the line connecting these 2 points is an integer point.

We shall consider two ways to solve this problem.

Proof. (Method 1). Assume the 5 integer points are $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, $P_3(x_3, y_3)$, $P_4(x_4, y_4)$, and $P_5(x_5, y_5)$. First consider their *x*-coordinates: x_1, x_2, x_3, x_4, and x_5. According to the Pigeonhole Principle, at least 3 of these 5 numbers have the same parity. Assume these three numbers are x_1, x_2, and x_3.

Consider the *y*-coordinates of the 3 integer points $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, and $P_3(x_3, y_3)$: y_1, y_2, and y_3. According to the Pigeonhole Principle, 2 of these 3 numbers have the same parity. Assume these numbers are y_1 and y_2. Now we have 2 integer points, $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$, where both x_1 and x_2 and y_1 and y_2 have the same parity. Therefore, both

$$\frac{x_1 + x_2}{2} \quad \text{and} \quad \frac{y_1 + y_2}{2}$$

are integers. This proves that the midpoint of the line connecting P_1 and P_2 is an integer point.

⑯ *Proof.* (Method 2). Use the following method to construct the holes. Consider the ordered integer pair (x, y). A total of four odd-even cases are possible for the ordered integer pair: (odd, odd), (odd, even), (even, even), and (even, odd). We construct 4 holes and put all 5 integer points into these 4 holes. Therefore, according to the Pigeonhole Principle, there is at least one hole that contains 2 integer points, say $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$. Therefore, both x_1 and x_2 and y_1 and y_2 have the same parity. This proves that the midpoint of the line connecting P_1 and P_2 is an integer point.

> In a three-dimensional plane, pick any 9 integer points. Prove that there must exist 2 of these integer points such that the midpoint of the line connecting these 2 integer points is also an integer point.

We can use the first proof method of the previous problem to solve this example, but this requires us to use the Pigeonhole Principle three times for the three pairs of numbers (x-coordinate, y-coordinate, and z-coordinate). This is a bit tedious, so we shall use the second proof method of the previous example.

Proof. Assume $P(x, y, z)$ is an integer point in a three-dimensional plane. There are total of

$$2 \times 2 \times 2 = 2^3$$

different kinds of even-odd combinations of the ordered triple (x, y, z). Therefore, there are 8 holes, and we have 9 integer points. Again, according to the Pigeonhole Principle, there is at least one hole that contains 2 integer points, say $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$. Then both x-coordinates, x_1 and x_2, have the same parity. Similarly, it is true for y-coordinates y_1 and y_2 and z-coordinates z_1 and z_2. This proves that the midpoint of the line connecting P_1 and P_2 is an integer point.

Prove: For a set of 27 different odd numbers, each of which is less than 100, there must exist 2 numbers whose sum is 102. NUMB

Proof. There are 50 odd numbers less than 100: 1, 3, 5, . . . , 49, 51, 53, . . . , 97, 99. ⑯ Within these 50 numbers, there are only the following 24 pairs of numbers whose sum is 102: $\{3, 99\}$, $\{5, 97\}$, . . . , $\{49, 53\}$; that is, $\{2k - 1, 103 - 2k\}$ $(k = 2, 3, \ldots, 25)$.

Consider the following number sets as 26 holes: $A_1 = \{1\}$, $A_2 = \{3, 99\}$, $A_3 = \{5, 97\}, \ldots, A_k = \{2k - 1, 103 - 2k\}, \ldots, A_{25} = \{49, 53\}$, $A_{26} = \{51\}$. (N.B. A_1 and A_{26} must be included to provide for the possibility of having the numbers 1 and 51 selected, although they cannot fit the program successfully.) And now we put these 27 odd numbers (less than 100) into these 26 holes. Then, according to the Pigeonhole Principle, there is at least one hole that contains 2 numbers. Since the sum of 2 numbers in every hole is 102, we have what we want to prove.

Randomly pick 50 numbers from the natural number set $\{1, 2, 3, \ldots, 100\}$ and prove that there must exist 2 numbers (of 50) such that one is the multiple of the other.

Proof. From the above several examples we should consider the following. In order to prove that there are 2 of n objects that satisfy the requirement, we need to construct $n - 1$ holes. For this problem we need to construct 50 holes, and every 2 numbers in the same hole have the property that one is the multiple of the other.

First, consider the 50 odd numbers from the natural number set $\{1, 2, 3, \ldots, 100\}$: $1, 3, 5, \ldots, 99$. Clearly, any number in the natural number set $\{1, 2, 3, \ldots, 100\}$ can be represented as the product of one number from the odd number set $\{1, 3, 5, \ldots, 99\}$ and some power of 2. In other words, we can separate the natural number set $\{1, 2, 3, \ldots, 100\}$ into the following 50 groups:

$$\begin{aligned}
&A_1 = \{1, 1 \times 2, 1 \times 2^2, 1 \times 2^3, 1 \times 2^4, 1 \times 2^5, 1 \times 2^6\},\\
&A_2 = \{3, 3 \times 2, 3 \times 2^2, 3 \times 2^3, 3 \times 2^4, 3 \times 2^5\},\\
&A_3 = \{5, 5 \times 2, 5 \times 2^2, 5 \times 2^3, 5 \times 2^4\},\\
&A_4 = \{7, 7 \times 2, 7 \times 2^2, 7 \times 2^3\},\\
&\ldots\\
&A_{25} = \{49, 49 \times 2\},\\
&A_{26} = \{51\},\\
&A_{27} = \{53\},\\
&\ldots\\
&A_{50} = \{99\}.
\end{aligned}$$

Every number in the natural number set $\{1, 2, 3, \ldots, 100\}$ belongs to one of the 50 groups. We can consider every group as a hole and then randomly pick any 51 numbers from the natural number set $\{1, 2, 3, \ldots, 100\}$ and put them into these 50 holes. Therefore, there must be at least one hole that contains 2 numbers. For any 2 numbers in every hole the bigger number must be a multiple of the smaller number.

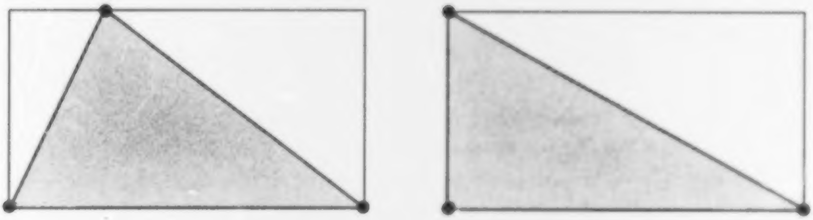
GEOM

There are 9 points in a square whose sides have length 1. Prove: For the triangles formed by these points, there must exist a triangle whose area is less than 1/8.

In this problem, what is going to be considered as a hole and how many holes exist is not yet clear. Therefore, first we need to construct holes. Before we do that we need to prove a corollary.

Corollary 1. For any three points in a rectangle with area a, prove that the area of the triangle formed be connecting these three points is less than a/2.

Proof. Because the largest triangle that can be drawn inside a given rectangle is one whose vertices lie on the sides of the rectangle (see Figure 8.1) and has an area half that of the rectangle, any triangle formed joining three points *inside the triangle* must have a *smaller* area than the previous triangle and thus *less than* half the area of the rectangle.



⑥

Figure 8.1

From Corollary 1 we know that for this example we only need to prove that there exists a rectangle of area 1/4 that contains at least three points.

Proof. Draw three lines parallel to one side of the square, and separate the square into four equal rectangles (Figure 8.2). Consider each small rectangle as a hole. Now we have 4 holes. Put these 9 points into these 4 holes. According to the Pigeonhole Principle, there is at least one hole that contains at least 3 points. From Corollary 1, the area of the triangle formed by connecting these three points is less than 1/8.

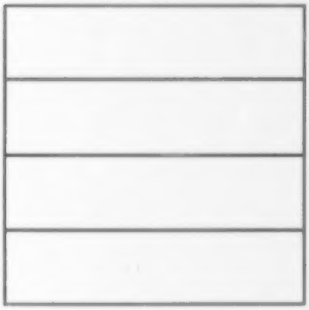


Figure 8.2

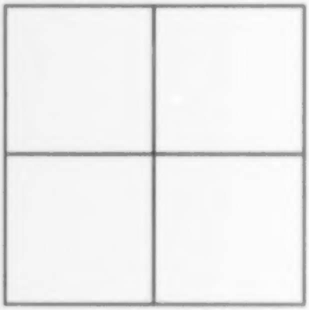


Figure 8.3

As an alternative method for solving this problem, two perpendicular lines are used to cut the square into four smaller squares each with area 1/4 (Figure 8.3).

There are 5 points in an equilateral triangle whose sides have length 1. Prove that there are at least 2 points separated by a distance of less than 1/2.

Proof. We may use the same method we used in the previous example. Connect the midpoints of the sides of the equilateral triangle. The original equilateral triangle has then been separated into four equilateral triangles (Figure 8.4). Let's consider each equilateral triangle as a hole. Then, according to the Pigeonhole Principle, there are at least 2 points in one of the holes. Since we know that all 5 points are in the original equi- lateral triangle, and the length of sides in the small equilateral is 1/2, the distance between the 2 points in the same hole is less than 1/2. This proves that, given 5 points, there are at least 2 of these points whose distance apart is less than 1/2.

⑥

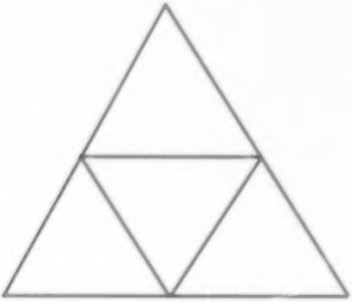


Figure 8.4

> There are 15 seats around a round table, and there are 15 customers' names posted around the table. Assume that no customers notice the name signs until they sit around the table and realize that no one is sitting in front of his or her own name sign. Prove that we can rotate the table so at least 2 customers are sitting in front of their own name signs.

GAME

Proof. We can separate the round table into 15 equal blocks and use $a_1, a_2, \ldots, a_{15}$ to represent the 15 customers' name signs (Figure 8.5a). Pick any customer from these 15 customers and use A to represent her. Assume that A initially sits in front of post a_1 (Figure 8.5b). From Figure 8.5b we should be able to establish the following: Whenever the table rotates one notch clockwise, there will be a name sign pointing to the seat where A sits.

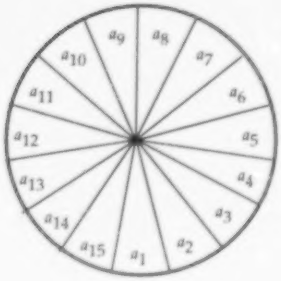


Figure 8.5a

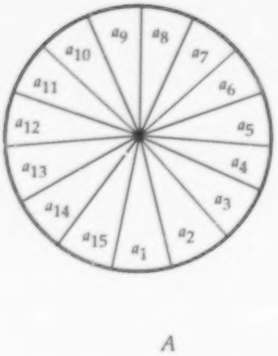


Figure 8.5b

⑥

After the round table has moved 14 notches clockwise, the 14 name signs on the table (except a_1) pointed at different times to the seat where A sits (Figure 8.6).

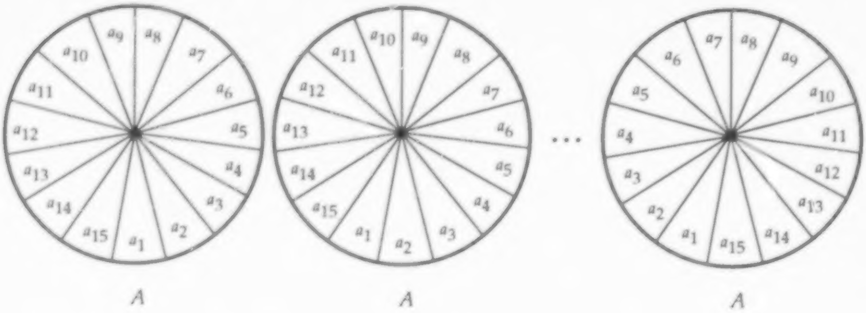


Figure 8.6

⑥

These rotation movements have all 14 name signs (except the one in front of which the person originally sat) point to every customer at the table, one at a time.

According to the initial assumption, no customer sits in front of his or her name sign. In other words, a_1 is not customer A's name sign. And customer A's name sign is among the other 14 name signs besides a_1. Then among these 14 rotations is one that has customer A sitting in front of her post.

This proves that every one of these 15 customers was situated in front of his or her name sign in one of these 14 rotations.

Now consider every one of these 14 rotation positions as holes, and the i^{th} ($i = 1, 2, \ldots, 14$) rotation position is called the "i^{th} rotation hole." If at the i^{th} rotation position customer A is sitting in front of her name sign, then put A into the i^{th} rotation hole. According to the Pigeonhole Principle, there are at least 2 customers in the same hole. In other words, there must exist at least one rotation position such that 2 customers are sitting in front of their respective name signs.

Part 1. There is a 4 × 7 block checkerboard (Figure 8.7), and every block is colored either black or white. Prove that in whichever way you color the checkerboard, it contains a rectangle consisting of more than one row *and* more than one column whose four corners have the same color (see one such example in Figure 8.7).

Part 2. Find a black-white coloring for a 4 × 6 block checkerboard so that there is a rectangle whose four corners do not have the same colors.

⑥

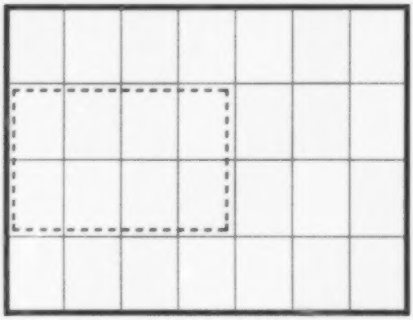


Figure 8.7

Before we solve this problem, let's analyze the question: According to the coloring requirements, every block must be colored by exactly one of the colors, black or white.

Figure 8.7 shows a checkerboard with 4 rows and 7 columns. The dashed ⑯ outline delineates a rectangle having all four corners the same color. The four like-colored blocks (in this case, white) shown in Figure 8.7 are positioned in the second and third rows and the first and fourth columns. The dashed rectangle may be called a "same-color-corner rectangle."

Part 1

Proof. (Method 1). Whichever way we color the checkerboard, we can always find the desired rectangle considering only three rows on the checkerboard—that is, a 3 × 7 checkerboard.

We shall first consider the first row. When we color the 7 blocks in either black or white, then according to the Pigeonhole Principle there are at least 4 blocks having the same color. Let's assume that these 4 blocks are colored white. Without loss of generality, we can assume that these 4 white blocks are positioned at the left side of the first row (Figure 8.8).

Now let's consider the other 2 rows. Since we are interested in the same-color blocks, we will consider only the first, second, third, and fourth columns and not consider the fifth, sixth, and seventh. If there are at least 2 white blocks in one of the other 2 rows, then these 2 white blocks, together with the 2 corresponding blocks in the first row, will be the four corners of the rectangle. This is a white same-color-corner rectangle. If every row has at most 1 white block, then delete the column(s) where they have the white block(s). The remaining 2 columns will form a black same-color-corner rectangle. ⑩

W	W	W	W			

⑥

Figure 8.8

Proof. (Method 2). There are 21 blocks in the 3 × 7 checkerboard (Figure 8.8). Consider the black blocks as black holes and the white blocks as white holes. Therefore, according to the Pigeonhole Principle, there are at least 11 blocks with the same color (say, white).

Now, let's consider every row as a row hole. Put these 11 white blocks into these 3 row holes. There are at least 4 white blocks in one of the row holes. In other words, at least one row contains 4 white blocks. Assume these 4 white blocks are positioned in the front of the i^{th} row. Now, as in the previous proof (method 1), we can conclude that there must exist a same-color-corner rectangle. ⑩

Part 2

We begin by coloring 2 of the 4 blocks in white in each column, and the other two blocks will then be colored black. There are $\binom{4}{2} = 6$ different colorings possible. This applies exactly to the 6 columns. Figure 8.9 is one such example, where there is no same-color-corner rectangle there.

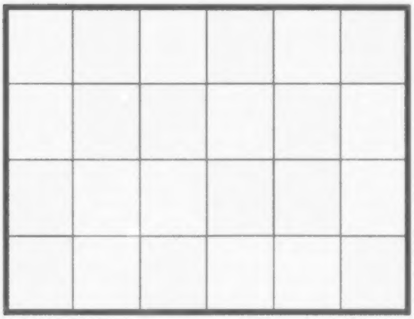


Figure 8.9

Exercises

1. Put any 6 integers into the 6 blocks (Figure 8.10). Prove that there exists a rectangle for which the sum of the four corner numbers is an even number.

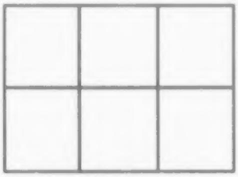


Figure 8.10

2. Given $n + 1$ points on a line segment of length 1, prove that there exist two points such that the distance between which is less then $1/n$.
3. At a fruit market there are 3 different kinds of apples mixed in one container. A customer wants to buy 3 apples of the same kind. The merchant picks 7 apples from the container, and it turns out that among them were 3 apples of the same kind. Use the Pigeonhole Principle to explain the reason why the owner picked 7 apples.
4. There are 50 students in a class. The oldest student is 18 years old and the youngest is 15. Prove that there are at least 2 students in that class who were born in the same year and month.
5. There are 7 points in a circle with radius 1. Prove that there are at least 2 of these points whose distance apart is less than 1.
6. Given 13 points in a rectangle whose sides have length 3 and 2, prove that within the triangles formed by any 3 of these 13 points there exists a triangle whose area is not more than 1/2.

7. Assume that everyone in New York City has at least 1 friend. Prove that there are at least 2 people in New York City who have the same number of friends.
8. Select any $n + 1$ numbers from the integers 1 through $2n$. Prove that there exist two of these $n + 1$ numbers whose difference is n.
9. If there are more leaves than trees, then there are at least 2 trees with the same number of leaves. Is this true?
10. Choose any $n + 2$ numbers from the integers 1 through $3n$. Prove that when $n > 1$, there must exist 2 numbers whose difference is between n and $2n$.
11. On a 3 × 9 block checkerboard, color the blocks red and blue. Prove that, whichever way you color them, there must exist 2 columns that have the same coloring.
12. Using 4 colors, color a 5 × 41 block checkerboard. Prove that, whichever way you color the blocks, there exists at least one same-color-corner rectangle.
13. Color the blocks of a 5 × 5 block checkerboard red and blue. Prove that there exists at least one same-color-corner rectangle.
14. Separate the positive integers from 1 through 67 into 4 parts. Prove that there must exist a positive integer in one part that is equal to the difference of two other numbers in the same part.
15. Select any $n + 1$ numbers from the positive integers 1 through $2n$. Prove that there must exist a pair of numbers in the selected group such that one number is the multiple of the other.

9

Handling, Seeing, and Thinking Experiences in Mathematics

EVAN M. MALETSKY

One of the great contributions of the early Greek mathematicians around the time of Pythagoras, Plato, and Euclid was to move an imperfect, hands-on mathematics based on experience and induction to a perfect and ideal, abstract mathematics built on deductive reasoning. They were very effective in establishing this view, and it has remained essentially unchanged to this day. Unfortunately, this unique mathematical strength may also be the cause for much of our current concern in mathematics education.

The National Council of Teachers of Mathematics (NCTM), in its *Agenda for Action* in the 1980s, identified the primary focus of the mathematics curriculum in the schools to be problem solving. Its *Teaching and Evaluation Standards for School Mathematics* of the 1990s broadly expanded this same theme by adding reasoning, communication, and connections. These recommendations give direction for the ideal experiences in the mathematics classroom, and they clearly incorporate handling and seeing as well as thinking.

Unfortunately, many believe these three terms form an ordered sequence.

handling → seeing → thinking

It is a fundamental premise here that real problem solving at all levels requires the freedom to move back and forth among the three.

handling ↔ seeing ↔ thinking

Furthermore, it is seeing that serves as the vital link between handling and thinking. If we are to enhance the problem-solving skills of our students, one thing we must improve is their ability to see properties, patterns, and processes, whether they be in numerical, algebraic, or geometric form.

⑥ GEOM

How often we assume, without question or doubt, that our students see what we see. How often we assume that because the numbers are there in a pattern to be seen, they see the pattern; that because all the words are there to be read, they see the problem; that because all the parts are there in a diagram to be seen, they see the whole of it. Many years of experience in the mathematics classroom have revealed an increasing deficiency in our students' ability to see and think in a visual, geometric form. It is our responsibility as mathematics teachers, regularly and repeatedly, to express numerical and algebraic relationships geometrically. Not only will this develop a better understanding of how these different fibers are interwoven into the fabric of mathematics, but it will also make geometric visualization a familiar and ready tool for problem solving.

Benoit Mandelbrot, considered by many to be the father of fractal geometry, writes: "Today, I believe deeply that science could accommodate more variety, and badly needs more people with a sharp and happy roving eye."

This is no idle wish. It is an apparent and frightening observation and should be a major concern to teachers of mathematics, especially when one considers the role the sense of sight can play in problem solving. The dilemma can be simply stated: Too many students look but do not see.

Looking Without Seeing

Do we train our students to look for answers with their eyes? And if we do, do we also encourage them to rove around and find what else they can see? Students learn from the way they are taught. They also see by the way they are shown to look.

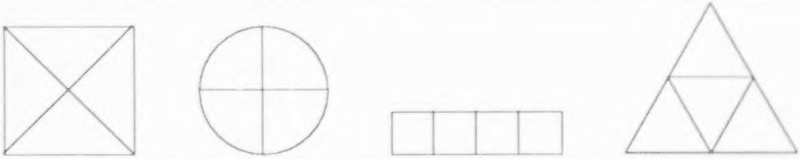
ARITH

Perhaps most important of all is the ability to see numbers in geometry and geometry in numbers. As a case in point, consider the following question from a recent National Assessment of Educational Progress (NAEP) test at the junior high level.

What is 75% of 12?

Surely, percentages remain one of the difficult topics to master at this level, but should this question, with these friendly numbers, be missed by more than half those trying to answer it correctly? Could it be that they looked but saw nothing? Could it be that the various algorithms used to teach percentage problems of this type did not include any that were visual? Perhaps this is why so many students failed to recognize these numbers as being friendly and, hence, could not associate the problem or its solution with any simple geometric figures.

The figures below all connect to the percentage problem above. It is the responsibility of the teacher, especially in the middle-school years, to show this connection not only through pictures to see but through pattern block, tangram, and paper-folding activities where things can be handled as well.



GEOM

Figures 9.1 through 9.4

Not only should the connection to geometry be made when doing friendly computation problems with percentages, fractions, decimals, and whole numbers, but the connection should also be made to these computations when doing the related work in geometry. Furthermore, this visual connection helps focus in on the fact that both the 75% and the 12 are friendly numbers for computation. One wonders how many of those taking the NAEP test even noticed this fact. We need to look carefully at how we teach what we teach and how we help students see when they look. Could it be that many students would see no difference between these two percentage problems and how to solve them?

What is 75% of 12?
What is 74% of 13?

One of the great skills in problem solving is the ability to look for differences as well as similarities and to know what to do with them when they are found.

The following problem appeared as a calendar problem in the NCTM *Mathematics Teacher*. It is used here to illustrate how we can enhance our students' ability to see numbers and patterns in what appear to be straightforward problems in geometry.

GEOM

⑥ Count all the triangles in this figure.

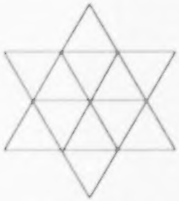


Figure 9.5

⑪ Two different skills are required here. One is the ability to see that there are three different sizes of triangles. The other is the ability to count systematically all the possible positions for triangles of each size. Cutting out a triangle of each size and moving it around on the figure can be a helpful strategy here. Some triangles point up while others point down. It may be less obvious that there are the same number of each kind in each size, but the horizontal axis of symmetry forces this to be true.

Table 9.1

Triangles	*Up*	*Down*	*Total*
1 - 1 - 1	6	6	12
2 - 2 - 2	3	3	6
3 - 3 - 3	1	1	2
Total of all triangles of all sizes			20

What else is there to see and do here? These follow-up questions illustrate just how many additional problem-solving strategies can easily be brought into play from this single problem.

What fraction of the area of one large triangle is shared with the other?
If a triangle is chosen at random, what is the chance that it is a small one?
What is the relationship among the perimeters of different-sized triangles?
How are the areas of the different-sized triangles related?
Do all pairs of middle-sized triangles share the same percentage of their areas?
Does the figure contain more triangles or trapezoids?
Are there any special number patterns in the figure?

Pattern blocks can be very useful in investigating number patterns in figures of this type. A part of the original figure is shown here as it might be built using the triangular pieces in a set of pattern blocks. In the figure on the left, the entire triangle is built with blocks. On the right, only those blocks that point up are used. ②

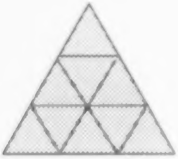


Figure 9.6

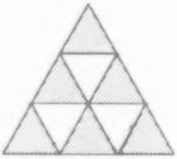


Figure 9.7

Have students search for geometric patterns. Then have them set up tables and look for numerical patterns that relate the number of blocks in given rows of the figures. Finally, have them extend the patterns for more rows. On the left the odd numbers are generated. On the right the counting numbers appear. ⑥

Table 9.2

Row	1	2	3	4	5	1	2	3	4	5
Number of blocks	1	3	5	7	9	1	2	3	4	5

NUMB

Have students look again for patterns, this time among the numbers giving totals from the top through a given row. Surprisingly, the perfect squares are generated on the left while the triangular numbers are generated on the right.

Table 9.3

Through row	1	2	3	4	5	1	2	3	4	5
Number of blocks	1	4	9	16	25	1	3	6	10	15

Once the separate number patterns are explored, let students investigate how they relate to one another as the sequence is extended. For example, as more and more rows are added to the array, what happens to the ratio of triangular holes to triangular blocks? Does it remain the same, increase, or decrease? If it changes, is there a limiting value approached as the number of rows increases without bound? ⑨

Table 9.4

Number of rows	1	2	3	4	5	6
Holes	0	1	3	6	10	15
Blocks	1	3	6	10	15	21
Ratio of holes to blocks	$\frac{0}{1}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{3}{5}$	$\frac{2}{3}$	$\frac{5}{7}$

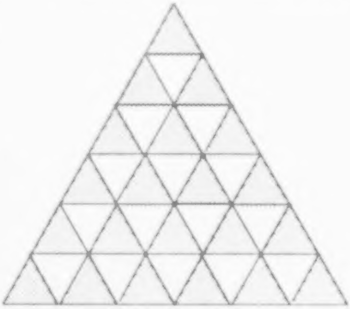


Figure 9.8

As the array gets larger and larger, the ratio of triangular holes to blocks increases toward 1.

Note how the dynamics of mathematics have been built into this last activity by looking not only at specific ratios but at how these ratios change. This sequencing of stages is the same kind of dynamics that can be built into the six-pointed figure from the original counting problem. Is it a static figure or is there some action involved? The answer, of course, depends on what you see. So let us look at it again.

Do you see the hexagon? Do you see the six triangles that form the hexagon? Think of them as petals on the bud of a flower, about to open. Now look at the complete figure again. Can you see it now as a beautiful six-petaled flower that has opened from a hexagonal bud?

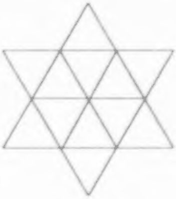


Figure 9.9

⑥ Challenge students to do the mental visualization first. Then have them cut the figure from paper, fold in the petals, and open it in their hands. Many students need to see the essence of a problem by drawing a picture or making a model. This flowering bud can serve as the basis around which some interesting questions can be built. Here are two examples.

GEOM

If the triangular petals may or may not open, how many flowers are possible in all, and in how many different shapes can they occur?

There are two choices for each of the six petals, open or closed. Thus, in all, there are $2^6 = 64$ possibilities. Visualizing the 13 possible shapes is a bit more challenging.

PROB

If there is only a 90% chance that a petal actually opens, what is the probability that the flower has the shape of a trapezoid?

First, see the six different ways a trapezoid can be formed from two petals open with one closed in between. Then multiply by the corresponding probabilities for the open and closed petals.

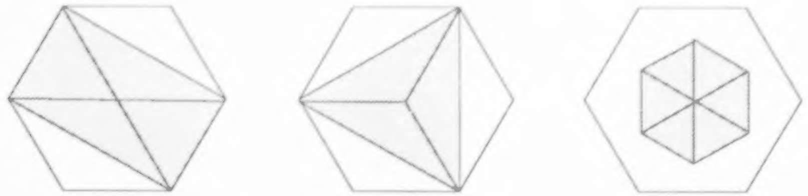
$$P\text{ (trapezoid)} = 6(0.9)^2\,(0.1)^4 = 0.000486$$

Good problem-solving experiences are those that require special, flexible, creative thinking. They often lead to new visions and frequently demand imaginative solutions. One way to help develop these skills is to encourage this same kind of play with all mathematical knowledge and experiences. ③

When one begins to sense the importance of handling, seeing, thinking, and connecting in a dynamic setting, it becomes easier to bring them into focus in even the simplest classroom situations. This activity with paper hexagons illustrates the point. It can help make mathematics come alive in the classroom and, at the same time, build skills that can be used later in solving other kinds of problems.

> Imagine some paper cut in the shape of regular hexagons. With one, fold an opposite pair of vertices to the center. With a second, fold alternate vertices to the center. With a third, fold all vertices to the center. What shapes do you get? What fractional part of the original area remains in each case?

Visualize the foldings first. What mental images do you see? Then fold with your hands and look with your eyes. Again, models can help in solving problems. Next, build a supporting argument for why the shapes are what they are. Finally, look at them numerically and find the resulting areas as fractional parts of the original hexagons. ⑥



Fold two opposite vertices
RECTANGLE
Area 2/3

Fold alternate vertices
TRIANGLE
Area 1/2

Fold all vertices
HEXAGON
Area 1/3

GEOM

ARITH

Figures 9.10 through 9.12

Clearly, challenging problem-solving experiences emerge from merely putting a simple paper hexagon into motion. Exploring these problems with the hands and eyes opens up new strategies for experimentation and observation that ultimately lead to solutions more rigorously established through deductive reasoning. ⑬

A Sample Problem

Counting activities from discrete mathematics offer rich sources of problems.

COMB A 9-inch piece of wire is bent at two points such that its ends come
GEOM together to form a triangle. If the bending points must be on the inch
marks, how many choices are possible?

Bring a piece of wire for the students to handle. Let one bend the wire so others can see a solution. Then let them think in groups about how to find all possibilities.

A typical strategy might be to find the different integer dimensions possible for the triangles by systematically accounting for all possibilities in a list.

Table 9.5

Length	*Classification*	
9-inch wire	Equilateral	3-3-3
	Isosceles	1-4-4
	Scalene	2-3-4

⑥

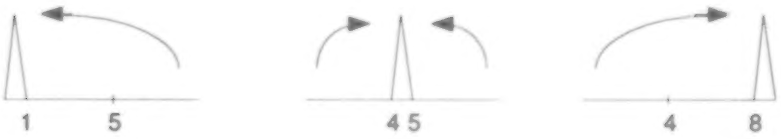
⑯ For some, the work ends here with an answer of 3. There are three different
③ triangles that can be formed. For others, a careful rereading of the problem reveals
the need for a second step that leads to an answer of 10 possibilities. The list shows
⑪ 10 different points where the wire can be bent in forming these triangles.

Table 9.6

⑭

Length	*Classification*		*Bending Points*		
9-inch wire	Equilateral	3-3-3	3 and 6		
	Isosceles	1-4-4	1 and 5	4 and 5	4 and 8
	Scalene	2-3-4	2 and 5	3 and 5	4 and 6
			2 and 6	3 and 7	4 and 7

⑦ Here are the three choices of isosceles triangles. In each case, the wire has been bent about the shortest 1-inch base, which has remained fixed in a horizontal position.



Figures 9.13 through 9.15

The first approach erroneously implies the three classifications are equally likely. The second approach shows that the chances are 3 in 10 that an isosceles triangle is formed, assuming, of course, that all 10 choices of bending points are equally likely.

As we as teachers do more with problem solving in the classroom, we become more comfortable with extending and adapting problems. Invariably, one good problem leads to another.

What if the initial wire were 10 inches long? Would there be fewer, the same number, or more folding choices for triangles?

The answer may come as a bit of a surprise. The wire is longer but there are fewer choices. The 9-inch wire gives 10 choices while the 10-inch length gives only 6 choices, all of which are isosceles.

Table 9.7

Length	*Classification*	*Bending points*		
10-inch wire	Isosceles 2-4-4	2 and 6	4 and 6	4 and 8
	Isosceles 3-3-4	3 and 6	3 and 7	4 and 7

Our goal as mathematics teachers is not only to improve our students' skills in problem solving but also to motivate and encourage them to search out, recognize, and explore problems on their own. The triangle problems with the 9- and 10-inch wires were only the appetizers. The main course comes when we look for an underlying rule for a wire of any integer length.

NUMB How and where does one begin? One approach is to guess and test some
different ideas. Another is to quickly solve some simpler analogous problems with
COMB very short lengths, such as with 3-, 4-, 5-, and 6-inch wires. The respective answers
of 1, 0, 3, and 1 may not reveal any clues. There are many different approaches to
try and many interesting connections to observe along the way.

⑦ Do you see anything special? For many, this means looking, in their minds,
at the triangles formed. Far fewer look at the answers of 10 and 6, along with the
1, 0, 3, and 1, and see if they can recognize or learn anything from them. Are these
④ special numbers? Is there a conjecture that you are willing to risk making and
checking?

The numbers 1, 3, 6, and 10 are all triangular numbers. Could there be any
② connection? Try some other cases and see. An 8-inch wire gives 3 choices, and an
11-inch wire gives 15 choices. That's both interesting and encouraging. They, too,
are triangular numbers, but the order in which they appear seems strange.

⑱ This is the place to send students off on their own for some more counting
and have them return (we hope) with an observation or maybe even a generaliza-
③ tion. This is also the place to encourage them to try adopting different points of
view, using numbers, algebra, or geometry with tables, formulas, or diagrams.

One possibility is to do some additional counting for other lengths to complete a sequence. Then collect the data and present them in the form of a table.
Here is a table of all choices for lengths from 3 through 15 inches. There is a striking
② pattern to be found, but not everyone will see it.

Table 9.8

⑨

Length	3	4	5	6	7	8	9	10	11	12	13	14	15
Choices	1	0	3	1	6	3	10	6	15	10	21	15	28

At all levels, pattern recognition remains among the most important of problem-solving skills, and yet it can often be very elusive. Just what is that special ability that lets some see the triangular numbers embedded once in the odd lengths and again in the even lengths, while others only see a single scrambled set of numbers?

This beautiful discovery only adds fuel to the mind of the critical thinker who finds in this solution both another good question and a chance to seek out another problem-solving strategy to reveal yet another answer.

How are the triangular numbers connected to this triangle-folding problem?

The triangular numbers are sums of sets of successive counting numbers, always starting with 1.

1 = 1
3 = 1 + 2
6 = 1 + 2 + 3
10 = 1 + 2 + 3 + 4
15 = 1 + 2 + 3 + 4 + 5

Had we approached the initial listing process from a different point of view, ③
we might have more quickly seen the relationship to the triangular numbers.
Instead of listing by classification according to sides, this method exhausts all those
with an initial bending point before moving to the next possibility. Notice how the ⑪
10 choices for the 9-inch wire are readily seen as the sum of 1, 2, 3, and 4
possibilities.

Bend at points 1 and 5.	1, 5			
Bend at points 2 and 5 or 6.	2, 5	2, 6		
Bend at points 3 and 5 or 6 or 7.	3, 5	3, 6	3, 7	
Bend at points 4 and 5 or 6 or 7 or 8.	4, 5	4, 6	4, 7	4, 8

Seen in a table of bending points, this triangular array becomes strikingly ⑥
visible.

	1	2	3	4
5	X	X	X	X
6		X	X	X
7			X	X
8				X

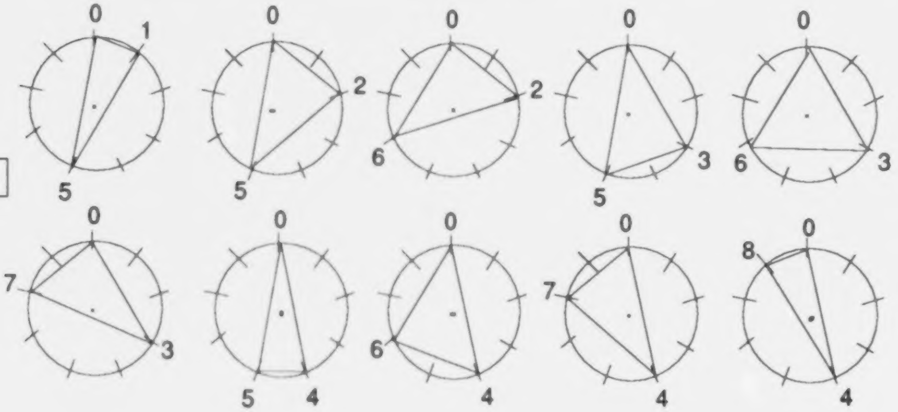
⑭

A Geometric View

⑥ The study of graph theory can help students view situations in geometric form. Graphs are pictures that offer a highly visual approach to presenting information, and they can be very useful in solving many problems.

③ Consider again the triangle-folding problem with the 9-inch wire. The three vertices of the triangle come from the two bending points on the wire plus the joining of the two endpoints into a single third vertex. The 10 choices of bending points are show here as connected graphs on circles divided into 9 units corresponding to the 9-inch length of wire. The three arc lengths on each circle correspond to what the actual lengths of the three sides would be.

GEOM



Figures 9.16 through 9.25

Why look at this problem in another way with pictures when the general pattern has already been discovered? The reason is that a different point of view often sheds new insight into a problem. It was the great mathematician and educator George Polya, who has been quoted as saying, "It is far more important to do one problem five different ways than to do five problems all one way."

As we look at the essence of problem solving in the mathematics classroom, it is wise to keep this idea in mind. Be careful not to measure success only by how many problems are done, but measure it also by what is done with those we do.

Consider again the 10 graphs showing the solutions for the 9-inch wire. What separates these from any choice of 3 of the 9 points on the circle? First, we use the point 0 to note where the two ends of the wire come together. The wording of the problem specifies choosing only two bending points on the length of the wire.

⑪ Combinations can be used to count the 28 possible choices of number triples that contain 0 as one of the choices. To do this, count the number of pairs that can be chosen from the remaining 8 points that can be connect with the point 0.

$$\binom{8}{2} = {}_8C_2 = \frac{8 \times 7}{2 \times 1} = 28$$

COMB

⑯ What distinguishes the 10 solutions from these 28 possible triples? Look at their graphs again, carefully. What do your eyes see? The sense of sight can be so important in the problem-solving process. We look, but what do we see?

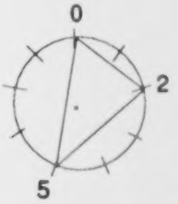
Some students' eyes only see the metric, measurement aspects of geometric figures. Others see graphs only in terms of points on a coordinate system. We seek different views here. The eyes may first spot that the center point of each circle lies within the triangle. Seen through another set of glasses, all these triangles are acute. What is the connection between these two observations, and how do they relate to the original problem?

⑬ Deductive reasoning comes into play here, recalling that the sum of the lengths of any two sides of the triangle must be greater than that of the third side. This means that the two bending points cannot be on the same side of the midpoint or at the midpoint of the wire, since the midpoint is not a unit point. Translated into our graphs drawn on circles, this requires the center of the circle to be enclosed within the triangle; hence, the triangle on the circle must be acute. Right triangles on the circle would contain the center point, since the hypotenuse would necessarily be a diameter. Obtuse triangles on the circle would leave the center point outside.

⑥ Does this mean that you can never form a right or obtuse triangle by bending a wire at the inch marks? The answer, of course, is no. Like many diagrams in discrete mathematics, our graphs on circles show connections and relationships in a discrete sense. The triangles shown on the circles are not the actual triangles folded from the wire.

For example, bending points at 2 and 5 form a triangle with sides of 2, 3, and 4 inches. The graph showing this choice of points may appear to be isosceles, but the actual triangle is scalene, since no two sides are the same length.

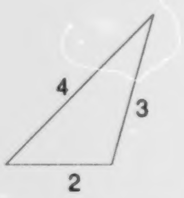
Graph



Bending points at 2 and 5.

Figure 9.26

Actual Triangle



Sides of 2, 3, and 4 inches.

Figure 9.27

The shortest wire length for a right triangle is 12 inches. Six of the 10 choices are scalene right triangles formed from the Pythagorean triple 3-4-5.

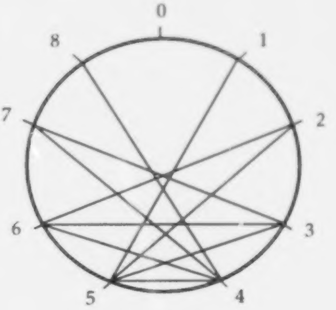
Table 9.9

Length	*Classification*	*Bending Points*		
12-inch wire	Right triangles 3-4-5	3 and 7	4 and 7	5 and 8
		3 and 8	4 and 9	5 and 9

The reader is asked to find the shortest integer length needed for an obtuse triangle.

⑥ As a final view of the problem, consider drawing all the triangular-connected graphs on the same circle. Better yet, since all triangles must contain a vertex at 0, only connect the remaining two vertices from each of the 10 graphs. Each chord now identifies a unique choice of bending points on the 9-inch wire.

9-inch wire



1 choice from point 1.
2 choices from point 2.
3 choices from point 3.
4 choices from point 4.

1 + 2 + 3 + 4 = 10 choices in all.

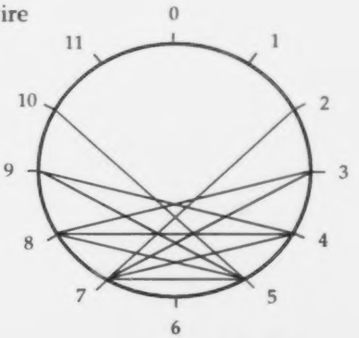
Figure 9.28

② This graph can give us insight into the connection between the triangular number pattern for odd lengths and the one for even lengths.

The even length that also has 10 choices is 12 inches. When the wire length
is odd, the first possible bending point is at point 1. But when the wire length is
even, the first bending point is at point 2. Furthermore, when the length is even,
⑬ the midpoint of the wire is a marked point but cannot be used. So 3 of the 12 points
on the circle cannot be used: the first and the last points (1 and 11) and the midpoint
(6). Thus the counting problem for the 12-inch wire reduces to that for the 9-inch
wire where all 9 points can be used.

Notice that the graph for the 12-inch wire has the same characteristics and the same number of solutions as the one shown above for the 9-inch wire.

12-inch wire



1 choice from point 2.
2 choices from point 3.
3 choices from point 4.
4 choices from point 5.

1 + 2 + 3 + 4 = 10 choices in all.

Figure 9.29

In each case, even lengths yield the same results as odd lengths 3 inches shorter.

Table 9.10

3-inch and 6-inch wires	1 choice	1
5-inch and 8-inch wires	3 choices	1 + 2
7-inch and 10-inch wires	6 choices	1 + 2 + 3
9-inch and 12-inch wires	10 choices	1 + 2 + 3 + 4
11-inch and 14-inch wires	15 choices	1 + 2 + 3 + 4 + 5

The diagrams illustrate the connection between the two triangular number patterns for odd and even lengths of wire. Another approach might be to show algebraically the two corresponding formulas. They yield the correct numbers but obscure the connection both to the triangular numbers and to each other. ③ ⑱

The k^{th} triangular number is the sum of the first k positive integers.

$$1 + 2 + 3 + 4 + 5 + 6 + \ldots + k = k(k + 1)/2$$ ALG

Let the number of inches in the length of the wire be n.
Let the number of possible triangle foldings be N.

odd length Use $k = (n - 1)/2$. $N = (n^2 - 1)/8$
even length Use $k = (n - 4)/2$. $N = (n^2 - 6n + 8)/8$

The Importance of Manipulative Experiences

Manipulative activities in the classroom can lead to rich problem-solving experiences when they tie handling and seeing to thinking. Bending a wire into a triangle may not appear at first to be a challenging activity, but we have just seen how fruitful an experience it can be.

Manipulatives quickly get students involved in the learning process. They enhance the chances for early success, which, in turn, paves the way for further, deeper investigation. The assumption that all students will be challenged by a good problem is simply false. Many need the encouragement of an initial success before they become willing or motivated to do some serious thinking at a higher level. The best problem-solving experiences for students are those that are just beyond their reach. Once they successfully stretch to reach a solution, their scope and vision of mathematics is likewise expanded and they are more likely to stretch to extend their reach yet again.

⑥ Manipulatives offer hands-on, visible objects that later become visual images for mental thought. They open up a whole new realm of geometric thinking, a skill that many problem-solving strategies require.

Manipulatives also help give students a dynamic view of mathematics. They encourage a sense of action, motion, and change—key components in the problem-solving process, where flexible views and alternate approaches must be freely tapped and tried.

The best manipulatives to use as settings for a problem-solving experience are usually the simplest. In this example, they serve to verify and reinforce as well as to explore and discover.

GEOM

A triangle is cut at random from a piece of paper and a vertex folded to the midpoint of the opposite side. What figures can result, and what determines which one appears?

⑥ ⑦ Have students think first and write down what they visualize as possibilities. Let then share and discuss their views and conjectures next. Then let them cut and fold and explore with their hands.

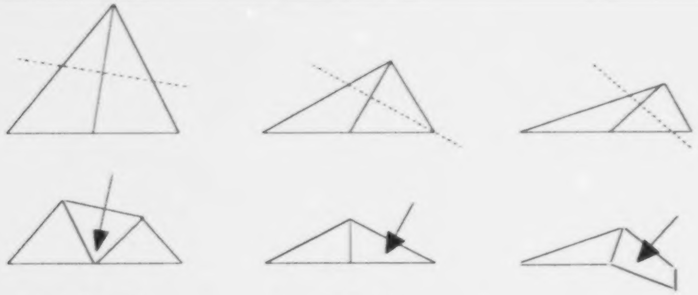
The initial challenge is to visualize in their minds the different possibilities. It is unlikely that many students will identify all three classifications of solutions. It many even be that they do not find them at the second step when cutting and folding with their hands. This step, with the actual manipulatives, offers insight and reinforcement and perhaps a surprise. But most of all, it motivates further investigation. Some students will study the shapes they get as results and look for the cause. Others will immediately start cutting and folding more triangles, hoping for an underlying rule to reveal itself to them. The reader is encouraged to explore and analyze this problem before continuing.

⑯

1. A quadrilateral is a rather obvious first possibility. The trapezoid is a special case when using the apex of an isosceles or any vertex of an equilateral triangle.
2. A pentagon is a second, less obvious, possibility. The fact that it is always concave may make it harder to visualize without actually folding in the hands.
3. A triangle is the elusive third possibility. It is also the starting point for an analysis of the entire activity.

Folding a vertex of a triangular piece of paper to the midpoint of the opposite side produces a crease that is the perpendicular bisector of the line segment connecting the vertex to the midpoint. Having the material in hand makes this observation considerably more apparent. ⑬

Notice the locations of the perpendicular bisectors in the three cases shown below. If the perpendicular bisector does not cut the opposite side containing the midpoint to which the vertex is folded, a quadrilateral is formed. If it intersects at either endpoint of the opposite side, a triangle is formed. However, if it cuts the opposite side into two segments, a pentagon is formed.



Quadrilateral Triangle Pentagon

Figures 9.30 through 9.35

For many students, this will appear as a rather abstract analysis. A more concrete observation, made from a hands-on exploration folding paper, deals with side lengths.

Let sides a and b contain the vertex to be folded with $a \le b$. Let c be the opposite side.

If $a > 1/2c$, then a quadrilateral is formed.
If $a = 1/2c$, then a triangle is formed. ⑱
If $a < 1/2c$, then a pentagon is formed.

It usually comes as a surprise that the result depends on the lengths of the sides of the triangles and not on the sizes of the angles. Each of the three cases of quadrilateral, triangle, and pentagon can be folded from an acute, right, or obtuse triangle. But, needless to say, different vertices may fold into different results.

Classroom experiences such as these are valuable because they connect different components of mathematics. But they are also valuable because they show students the importance of using their eyes in looking for solutions to problems. Even analytic solutions require visual skills, as in this next problem.

A triangle is cut at random from a piece of paper. What is the probability that it is an acute triangle?

GEOM
STAT
⑭

This certainly appears to be a hands-on activity. Just cut out a lot of triangles at random and see what proportion are acute. If the question is asking about actual results from actual cutting, collect some data and study the results. This is often a very powerful problem-solving strategy.

③

Alas, we so often require in mathematics a theoretical approach that is ideal and abstract, deductive rather than inductive. After all, how else can randomness really be considered without personal biases? However beautiful such solutions are, they remain, from one point of view, impractical. Yet they illustrate the essence of mathematical thinking so we offer one approach here. (There are others, some of which yield different results.)

⑬
ALG

Here is one analysis based on deductive reasoning involving algebraic relationships. Let the three angles of a randomly chosen triangle have degree measures of A, B, and C. That implies these conditions.

⑱

$$A + B + C = 180, \text{ with } A > 0, B > 0, \text{ and } C > 0$$

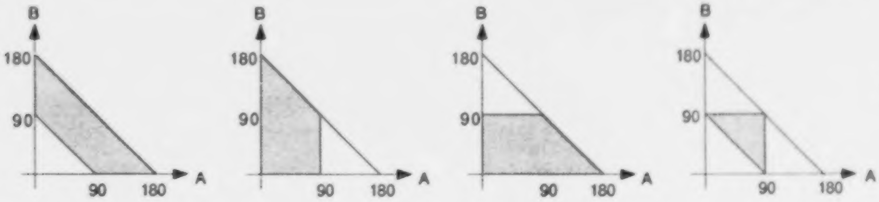
For the triangle to be acute, these additional conditions must hold.

$$A < 90, B < 90, \text{ and } C < 90.$$

Clearly, C is dependent on A and B. Therefore, we can restate the conditions in terms of A and B alone.

$$90 < A + B < 180, \text{ with } A < 90 \text{ and } B < 90$$

What do these inequalities look like when shown on a graph? The sample space is a large right triangle where every point in its interior corresponds to an ordered pair of values (A, B). Assuming all such points are equally likely, selected areas can be compared to find appropriate probabilities. The areas, associated with the three separate conditions, are shown first, with their intersection given at the far right.



$90 < A + B < 180$ $A < 90$ $B < 90$

⑥

Figures 9.36 through 9.39

Many students will be able to establish the necessary inequalities but may not know how to use them to get the needed probability. Others will draw the graphs but may not be able to translate the results into a probability because they do not see all that is there when they look at the graph. The four small triangular regions are all equal in area but only one of them contains points that satisfy the inequalities required for an acute triangle. It follows that the theoretical probability of randomly cutting an acute triangle under these assumptions is 1/4. In similar fashion, we conclude that the probability of cutting an obtuse triangle is 3/4. Now this very same sense of sight that was so useful thus far challenges us to accept 0 as the probability of randomly cutting a triangle that is a right triangle.

$$P\text{ (acute)} = 1/4 \qquad P\text{ (obtuse)} = 3/4 \qquad P\text{ (right)} = 0$$

PROB

It is interesting to note here that, were we to answer the problem by collecting data using triangles actually cut from paper with scissors, substantial numbers would likely say they cut right triangles.

Be cautious of the potential of a computer simulation here. It will give nothing better than an approximation to some ideal approach to handling the randomness question. In this attack, we assume triangles are chosen by randomly selecting the three angles. A computer simulation requires a program, and that program requires a definition of how the random triangles are chosen. If we choose this same approach, our computer results can do no better than approximate the answers given above. ⑫

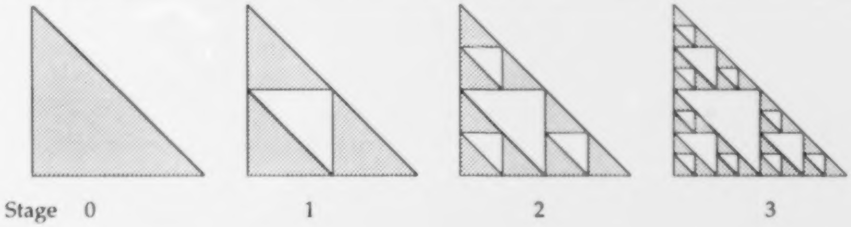
Seeing Things That Cannot Be Seen

The ability to visualize in the mind comes, in large measure, from the ability to see with the eyes. To a great extent, visualization is an extension of the sense of sight. There are many times in mathematics when perfect, abstract things are seen in the mind as extensions of imperfect, concrete things that we can handle with our hands and see with our eyes. The key components of Euclidean geometry—points, lines, and planes—are examples, as are circles and squares and spheres and cubes. Indeed, the very triangles that have been mentioned throughout this chapter are examples as well. We view them once again, but this time as part of a dynamical system. Students need experiences seeing changes that result from dynamical, iterative processes, especially since sequencing is, in fact, another problem-solving strategy.

> Start with a triangle cut from paper. Connect the midpoints of the triangle to form four smaller triangles. Keep the three corner triangles but discard the middle one. Then repeat the rule with every new triangle at each and every stage.

② The process produces a sequence of figures, growing successively ever more
⑨ complex. The initial triangular region is shown along with those that remain at successive stages. What patterns emerge as the iteration continues?

GEOM



Figures 9.40 through 9.43

Here are some counting, measuring, and geometric views of this sequence.

A Counting View

③ How is the number of triangular regions changing?
⑥ How is the number of triangular holes changing?

A tabular array can be used to investigate these relationships.

Table 9.11

NUMB

Stage	0	1	2	3	4	*Stage* n
Triangles	1	3	9	27	81	3^n
Holes	0	1	4	13	40	$\frac{3^n - 1}{2}$

At stage 5, the number of triangles is $3^5 = 243$. At stage 5, the number of holes is the sum of the zero through fourth power of 3.

$$3^0 + 3^1 + 3^2 + 3^3 + 3^4 = 1 + 3 + 9 + 27 + 81 = 121$$

The ratio can be computed at any stage, but another approach is required to find its limiting value. Each successive ratio can be shown to be closer and closer to the fixed point attractor of 1/2. ⑤

A Measurement View

How is the total area of the shaded regions changing? ③
How is the total perimeter of the shaded regions changing?

Again, a tabular array is used to investigate these relationships. Note that for convenience, an arbitrary unit of area and of perimeter has been assigned to the stage 0 figure.

Table 9.12

ARITH ⑥

Stage	0	1	2	3	4	*Stage* n
Area	1	3/4	9/16	27/64	81/256	$(3/4)^n$
Perimeter	1	3/2	9/4	27/8	81/16	$(3/2)^n$

At stage 5, the total area is $(3/4)^5 = 243/1024 = 0.237$. At stage 5, the total perimeter is $(3/2)^5 = 243/32 = 7.594$. ⑨

The dynamics of change are intriguing here. While the area is decreasing from stage to stage, the perimeter is increasing. As the iteration continues, ad infinitum, the area shrinks towards zero while the perimeter increases without bound. ⑤

A Geometric View

③ We can draw the first few stages with our hands. We can see and extend the geometric pattern with our eyes. We can devise a new three-step transformation algorithm for generation.

Reduce the figure to half size.
Replicate three times.
Rebuild in the original form.

But can we stretch our minds to visualize the strange attractor that is being approached by these successive stages?

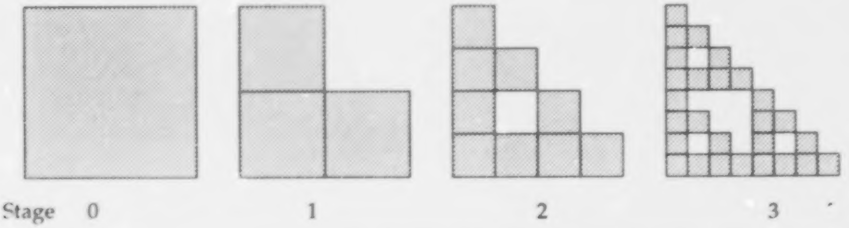
Here we must see with our minds what we cannot see with our eyes. As the iterative sequencing continues, more and more of the area of the original triangle is eaten away. Each and every triangular region at each and every stage is being reduced in successive stages. At the final limit state, the original triangular region has been completely eaten away with holes so no area at all remains. Each small ⑯ triangular region is ultimately reduced to a point with an area of 0. But there still is something there!

Infinitely many line segments remain, forming a delicate, endlessly repeating, ever-diminishing, highly complex fractal framework. To help in this visualization, notice that the segments that form the boundaries of the triangular regions at each stage remain as part of this final figure. This strange attractor, the resulting fractal, is called the Sierpinski triangle. It is named after the Polish mathematician, Waclaw Sierpinski, who discovered it in 1917.

One of the striking properties of the Sierpinski triangle is that it possesses strict self-similarity. Take any small piece of it, and embedded within that piece will be an exact but reduced replica of the original figure. Another property of this and all fractals is that its fractal dimension can be measured. The Sierpinski triangle is more complex that a line segment with dimension 1 and less complex than a region of the plane with dimension 2. Hence, we would expect its fractal dimension to lie between 1 and 2, as it does with a value of 1.58.

Aspects of fractal geometry may soon be in the secondary school curriculum. Not only is this topic accessible at these levels, but it also captures a new view of mathematics as dynamic, active, and alive. However, it requires substantial prerequisite experiences in the classroom, especially those that utilize the action and motion of manipulatives. It builds on the handling, seeing, and thinking activities of good problem-solving experiences of earlier years.

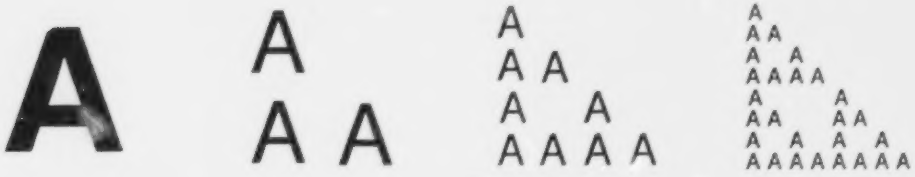
Fractal geometry offers students an excellent opportunity to search and rove with their eyes and to stretch and flex with their minds. To offer a final illustration of this point, study these results of an iterative process applied to squares. At each stage, every square region is replaced with three smaller ones such that the total area is repeatedly reduced to three fourths that of the preceding stage.



Figures 9.44 through 9.47

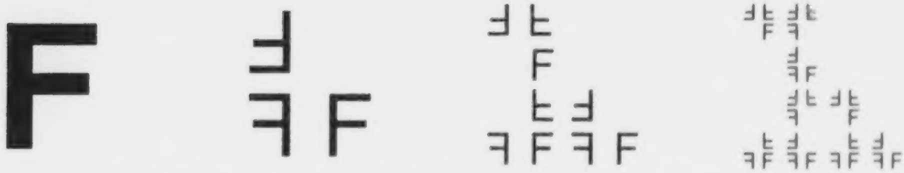
Each and every figure consists of shaded square regions, increasing in number and decreasing in area. But what is the strange attractor that is being approached here with the square regions in these successive stages? Surely, it is some self-similar fractal framework based on squares. As much as the eye tells us that it is, it is not. The resulting fractal is the same Sierpinski triangle described earlier.

When viewed as a transformation algorithm, this one with squares behaves exactly the same as did the one that started with triangles. The rebuilding code is shown in the squares. Whatever the figure inside the square, when reduced to half ② size, replicated three times, and rebuilt by that code, the resulting fractal will, ⑥ indeed, be the same Sierpinski triangle. The figure inside the square could be a triangle or any other shape whatever. Look with your eyes at the first few stages ⑨ using the letter *A* and this same rebuilding code. Then see if you can visualize the next few stages in your mind and recognize the strange attractor. ⑯



Figures 9.48 through 9.51

Keep in mind the way the rebuilding is repeated at each new step. After the reduction to half size and the threefold replication, the rebuilding code involves only translations of the squares containing the letters *A*. When rotations and reflections of these squares are involved in the building codes, new shapes are formed with new and different strange attractors. GEOM



Figures 9.52 through 9.55

How does one solve the problem of getting a glimpse of the general shape of these final fractals? We can resort to yet another problem-solving strategy, that of using a computer. With a computer, enough additional stages can be drawn so that the eye can begin to see and perceive in those images the characteristics of these fractals, the beautiful mathematical abstractions of the imagination, that are evolving.

⑫ The computer-generated figures below show higher stages in the sequences for the two square rebuilding codes that were used above. Already the general shapes of the final fractals are clearly emerging. However, the computer is limited to some small finite pixel size. Remember that it is only the final limit states of these figures that are the real self-similar fractals.

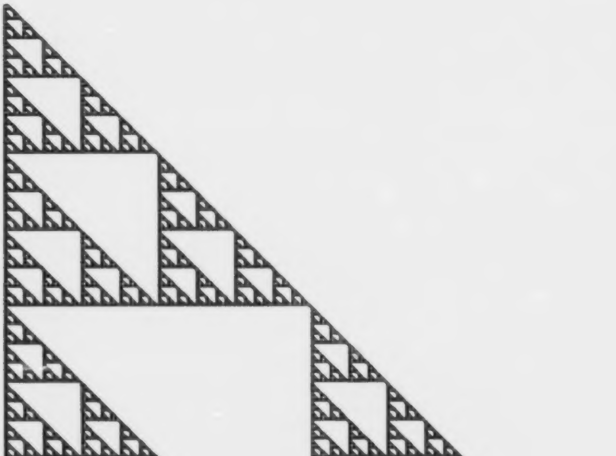


Figure 9.56

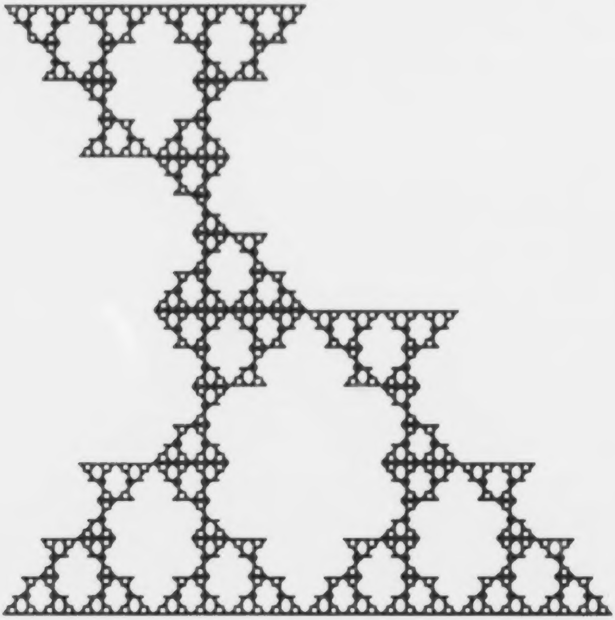


Figure 9.57

The real study of these types of fractals begins with their limit figures and their self-similarity. Hidden in them are the building codes. The real visual experience is in finding and identifying these transformations from their limit figures.

Fractals stretch the imagination because they are so highly visual in their repetitive, diminishing, complex patterns. Computers become useful tools to help us see into the depths of these figures and study the details that we cannot otherwise see with our eyes. The focus of this chapter has been on the sense of sight and its role in problem solving, and this visual experience with fractals stretches that experience to the limit.

Conclusion

Fractals offer excellent examples of what the NCTM means, in its *Curriculum and Evaluation Standards for School Mathematics*, by connections within mathematics. They also bridge the gap between mathematics and the sciences and the arts. Within the past several years, fractals have found applications in virtually every discipline. They offer a new view of the world around us, a model perhaps more

⑥ powerful and appropriate even than the Euclidean models of ancient times. But most of all, they develop two very critical skills in problem solving: insight into geometric thinking and visualization.

As we teach, we want our students to work with their hands, to see with their eyes, and to think with their minds. Our goal is to ultimately help them think critically and to solve problems, not just in mathematics but in all areas. We want them to apply what they know to real problems in everyday life dealing with real things that are often imperfect, approximate, and concrete rather than perfect, ideal, and abstract. We can help them achieve this goal though problem-solving experiences in the classroom that focus on handling, seeing, and thinking. These are what the mathematical experience is about.

Mathematics must tickle the senses as well as stretch the mind.

10

Problem Solving as a Continuous Principle for Teaching

Suggestions and Examples

HANS HUMENBERGER

HANS-CHRISTIAN REICHEL

1. Introduction

It is not the aim of this chapter to discuss problem solving as generally as possible or to deal with strategies of problem solving in a systematic way. Many authors already have dealt with all these topics in various ways (e.g., Dürschlag, 1983; Engel, 1979; Polya, 1954, 1980; Sell, 1988; Zimmermann, 1983).

The emphasis of this contribution is clearly on a more or less systematic presentation of selected examples (problems), including detailed discussions of their solutions. We would consider ourselves most successful if teachers were to make use of some of these problems in their classes or if they were stimulated to look for, or to create, similar problems themselves so that students can profit from a lot of (we hope positive) experience in solving such problems. In that way, we would make a small contribution to mathematics as well as to the teaching of mathematics.

We begin with some general introductory notes. Problem solving is an old topic in the teaching of mathematics, but it is not dealt with in a completely satisfying way. Many mathematicians and mathematics teachers dedicated their work to "problem solving" or to "drawing plausible conclusions," respectively. Many "strategies" among a large number of possibilities dealing with problem solving were (and are) emphasized as especially important; they are often demonstrated by some tasks, and sometimes students even get a "catalogue" on how to proceed and are then expected to find systematically the adequate strategies for solving the problems. Nevertheless, the following questions seem to us to remain not yet fully answered: Is it possible to teach heuristcs or problem solving systematically and to integrate it into the curriculum? Can a problem-solving process really be divided into "phases"? Is it possible to ensure a successful solution of a problem by a "strategy-pattern"?

We don't propose to give a definitive answer to these questions! We do not know if the "phases" of problem solving manifest themselves for everybody in the same way, if they appear in an ordered and comprehensible way at all, or if some phases appear parallel, disordered, and indescribable. A satisfying comprehension of a problem-solving process is hardly possible, because our thinking often appears in an un- or subconscious way; it is often not possible to say how one got certain ideas, or why one chose a certain way. In many cases one is working unfounded, trying and choosing ways without knowing where they will lead.

Despite all uncertainties about the systematics of problem solving, we do have the conviction that it is very important to provide enough opportunities in mathematics classes to solve problems that do not always fit into learned patterns, to be creative, to give reasons (to argue), to try (possibly wrong) ways, to get experiences, and so on. It seems a common practice among teachers to restrict problem solving to the better students while presenting students in the average compulsory classes with certain mathematical methods and later drilling them. We are not against "drilling certain methods"—quite the contrary—but we feel this should not be all of the classwork.

It is very important to find tasks that are "problems" in the sense that they promote creativity and spirit, that they motivate, that they are not limited to simple exercises but also are not too tricky (thus giving average students a realistic chance). Students can get self-confidence and be further motivated by having successful experiences. We are thinking of motivating and interesting tasks that contain "real mathematics" that fit into the curriculum, and that give students an appetite for more mathematics—not only gifted students but also for average and below average students!

We would not go so far as to suggest that all classwork should be based on problem solving, but we believe that one can (should) pose suitable and interesting problems for each mathematical topic, as "application tasks" or, later, as "repetition tasks," which show that the acquired mathematical knowledge was not just important "for the test." This should start with the beginning of mathematics classes and should recur as a continuous class principle. Furthermore, it seems important to us that the students get enough time to deal with the problems themselves and that various possibilities are discussed thoroughly and are really understood by most of the students. There must be time for mistrials as well, and these must not simply be considered as "wrong"; there has to be an explanation why certain methods do not lead further—the problem must really be "digested"! It is of only little use for students to be rushed from one problem to another, or to have the teacher distinguish himself or herself with brilliantly demonstrated solutions (which appear out of nowhere) and where the student cannot even guess where the statement came from or why the teacher tried one way and not another. Learning to solve problems surely does not depend on which problems the students have to solve, as much as on how the lessons are shaped, what activities are expected from the students, and what the climate is in the classroom.

But the treatment of problem solving as a "leading idea" for classwork is not without difficulties, and we hope our chapter and this book will alleviate some of these problems:

- Tasks (problems) with a suitable degree of difficulty that are interesting and motivating are often not sufficiently presented in textbooks, and it is often not easy for teachers to find or create such tasks. Therefore, this chapter will offer stimulation in the form of a variety of mostly annotated examples that are aimed not at performance-oriented math team groups but primarily at "normal" mathematics lessons in schools.
- The judgment of whether or not a task is "suitable," in the above-mentioned way, is naturally a subjective matter, so that there does not have to be agreement between the authors or editors of particular collections of examples, teachers, and students who may consider some suggested problems boring or too hard or not usable for other reasons.
- As we mentioned already, a lot of problems can be solved only with the help of some tricks, but to draw the line between a trick and an important mathematical method is not easy and is also very subjective.

To summarize, we want to stress once more that it is definitively not enough to demonstrate problem solving once or twice a year to the students or to show them how a teacher is easily able to solve even difficult problems. Students must have various opportunities to get down to easier problems themselves so that they get a chance to think about them and to develop solutions themselves (maybe with some minor hints). Even if these solutions are not always correct, students can develop confidence in their own abilities and lead them to find the courage and joy to do more problems or more mathematics in general.

The following tasks are not ordered by diverse heuristic strategies but by two different aspects. The first section will deal with problems in which a colloquial text has to be translated into the "language of mathematics" (in most cases this will mean an equation, but not always!). In our opinion, this type of exercise should be practiced particularly early and often. The translation of a colloquial problem into the language of mathematics is frequently "half of the solution" to a problem (cf. Reichel, 1991).

The second section consists of examples that clearly illustrate certain theories (or parts of theories) and help to make them more understandable. Moreover, they could promote motivation with their surprising results, and they might be well suited as introductory problems in certain fields. Even new mathematical topics can be introduced by simple problems, which show clearly the necessity and possibility of an extension of students' current knowledge with a generalization or a new theoretical concept. But it is exactly this type of problem that is often hard to find.

2. Translation of Texts Into the Language of Mathematics

One very important type of problem solving is translating colloquially described situations ("texts") into the language of mathematics. In such problems, it is first necessary to work out the mathematical structure of the task. One has to get an overview of the described situation, although it may not be immediately clear to show how it is possible to apply mathematics. It is especially important to deal with problems that allow more than one approach to the solution as well as problems that might contain superfluous data for students to detect. However, they should not be so complex that only extremely gifted students can manage them. The mathematical analysis of a problem will, in many cases, lead to an equation, but not necessarily (e.g., Example 2.3).

One wrong idea we want to confront decisively is that problem solving is interesting only for those students who have already achieved a relatively large, basic knowledge and who are at least 14 or 15 years old. It is also possible to formulate interesting problems for younger students (for instance, 10-year-olds) on a lower mathematical level. Simpler problems of this kind can be motivating for these students and can even help them in their further mathematical education. In the following sections, we want to present some elementary examples that can be used for students as young as 10.

2.1. *Simple Overdetermined Problems*

Working with over- and underdetermined problems is, in our opinion, a fine means to "sharpen the spirit," to promote the ability of the students to better recognize the structure (the essential things, the gist), and to translate (colloquially formulated) texts into questions that can be answered in a mathematical way. This capability is combined with the ability to apply mathematics—an ability everybody will need to solve his or her (mathematical) problems (in school and later, too) and which we should therefore master!

Example 2.1

The freshman class of a high school in a small town of 4,300 inhabitants takes a trip to a mountain 120 km away. There is $500 in the class treasury. The total cost of this trip was $360. This amount was needed to pay the bus fee of $110 and to pay the cost of the rope-walk for each of the 25 students.

ARITH

⑭

a. How much was the rope-walk for one student?
b. Which data were not necessary to answer question *a?*
   Rewrite the text using variables instead of specific numbers ($a, b, c, \ldots$) and give a general formula for the required fee.

Here the students first have to find out which data are not necessary (freshman class, 4,300 inhabitants, 120 km, $500). Then they have to recognize the mathematical structure of the task: The total spending consists of the fees for the bus and rope-walk. The students should be able to find the formula for the student's fee p basically by themselves: $p = (360 - 110)/25 = 10$.

Such examples can be embellished and extended almost limitlessly, and there are no bounds for the teacher's imagination. The more sophisticated and complex the story is and the more details (variables or concrete numbers) that appear in the text, the more difficult it is, in general, to find the unnecessary data, particularly if several quantities are asked for. (Many students can visualize a situation much better with concrete numbers.) The students could also be asked to develop such "task-stories" themselves (e.g., in smaller groups where students then solve each other's tasks or as homework.) This may be a small contribution to promoting the student's motivation, especially the very young student's (10 years old).

Analogous to this, tasks with too little information for the solution—"underdetermined" problems—would be valuable too. Here the students would have to figure out why the problem is not solvable, and they could make suggestions as to what additional information is needed to solve the problem.

Reality—in contrast to math textbooks—often does not provide the full amount of data needed to solve given questions. On the other hand, we often have much more information than necessary. Therefore, to educate in the spirit of applied mathematics, textbooks should copy reality in this respect (even artificially) so that students have to figure out which and how many data have to be known to answer the questions asked before they start working. Students should also be able to say explicitly which additional data would be necessary. Such problems are called over- and underdetermined problems in this chapter.

Example 2.2

ARITH

An employee is going to work by bicycle. Usually he goes the 3 km distance at an average speed of 15 km/h. But this time he was unlucky because, after 1 kilometer, he got a flat tire, so the journey took him 20 minutes more because he had to push his bike from this point. Fortunately, he was able to repair the damage at work and could cycle home as usual.

a. How many kilometers more did he go by bike than he walked?
b. Which data are irrelevant to answer question *a?*
c. Which additional questions could be asked and answered?
 or
d. Is it possible to answer the following questions with the given data?
 (i) How long does the bike trip normally take him?
 (ii) How long (how much time) did he have to push his bike?
 (iii) What was the average walking speed?
 (iv) How long did the repair take him?

Solution: A sketch showing the distances that were walked and cycled may be very helpful (Figure 10.1).

⑥

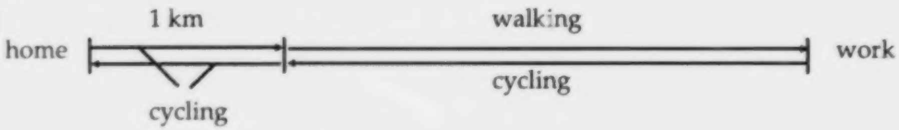


Figure 10.1 A cyclist who had some bad luck on his journey

To solve *(a)*: He had to push his bike from the scene of the accident to his work, but on his way back, he rode his bike exactly the same distance as he had to walk earlier (therefore the distance does not matter). The distance from home to the scene of the flat tire is covered twice by bike, once on the way out and once on the way back, so he went 2 × 1 km = 2 km more by bike than he walked.

⑭ To solve *(b)*: Everything is unnecessary except the 1 km! The value in question *(a)* does not even depend on the total distance of the trip.

To answer the four parts of *(d)*:

(i) $\frac{3}{15}h = 12$ min

(ii) 28 min

(iii) $4\frac{2}{7}$ km/h

(iv) not answerable

Example 2.3

ARITH

Mr. Mayer usually takes the 7:25 bus from home to his office, and in the afternoon he walks home (for health reasons). This takes him altogether 1 hour and 10 minutes. If he were walking both to and from the office, it would take him 1 hour and 50 minutes, because on average he walks only 15 km/h slower than the bus travels, with its scheduled stops, traffic lights, congestion, and so forth.

a. How long would it take him to go both directions by bus?
b. What data are unnecessary to answer question *(a)?*
c. What questions could be answered with the given information?
   or
d. Is it possible to answer the following questions with the given information?
   (i) When does Mr. Mayer arrive at his office?
   (ii) How fast does Mr. Mayer walk, on average?
   (iii) How fast does the bus travel, on average?
   (iv) What is the distance from Mr. Mayer's office to his home?

Solution: For *(a)*; As walking both ways takes him altogether 1 hour 50 min, walking one way takes him 110/2 = 55 min. Therefore, one way takes the bus 70 min – 55 min = 15 min. Hence, both directions would take the bus 30 min.

To solve *(b)*: The data 7:25 and 15 km/h are superfluous.
To answer the four parts of *(d)*:

(i) Normally he arrives at his office at 7:40.

(ii) $5\frac{5}{8}$ km/h

(iii) $20\frac{5}{8}$ km/h

(iv) $5\frac{5}{32}$ km.

The latter three solutions are to be expected at the earliest from 12- or 13-year-olds.

2.2 *Problem Solving by Equations—Text Equations*

ALG The following example will show, among other things, that it is possible, even at an elementary level, to demonstrate how important it is to mark the right quantities with variables.

Example 2.4

> Forty-two birds are sitting on three trees. If 3 birds fly from the first tree to the second one and 7 birds fly from the second tree to the third one, there are twice as many birds on the second tree as on the first one and twice as many on the third tree as on second one. How many birds were sitting originally on each tree?

Probably a student's most familiar method of translating texts into the language of mathematics—to define the original bird-numbers with x, y, and z —will lead to a system of three linear equations with three unknowns:

$$x + y + z = 42$$

$$2(x - 3) = y - 4$$

$$2(y - 4) = z + 7.$$

Their solutions would be $x = 9$, $y = 16$, and $z = 17$. But such equation systems are solvable only for students who are at least 14 or 15 years old (at least in Austria). So one could ask the students to look for a solution that does not involve solving a system of equations. Alternatively, if this problem is stated already for 12- or 13-year-old students, then they are forced to find a different method anyway.

① The situation, arithmetically, gets much easier if one looks at it first after the birds' flights: If there were a birds on the first tree, there were $2a$ birds on the second tree and $2\,(2a) = 4a$ birds on the third tree. That makes altogether $7a$ birds, whose total number is still 42. Therefore a—the number of birds on the first tree after the birds' flight—must be 6 ($42/7 = 6$). So after the birds' flight, there were 6, 12, and

24 birds on the trees 1, 2, and 3, respectively. Now we have to "calculate backwards": Since 3 birds flew from the first tree there had to be 6 + 3 = 9 birds in the beginning. An additional 7 birds came to the third tree, which means that initially there had to be 24 – 7 = 17 birds. Therefore, 16 birds must have been sitting originally on the second tree (42 – 9 – 17 = 16). One can do the following to check: 3 birds came to the second tree and 7 left it—which means that after the flight there had to be 4 birds less than before, and 12 + 4 = 16.

The second method does not contain any mathematics unknown to 12-year-old students, so it is, at least arithmetically, a little bit more economical. But in this case, one has to recognize first that the simple situation of the birds after the flight makes a better basis for the solution of the problem.

Example 2.5

> Two athletes—Anton and Ben—start running from opposite ends (*A* and *B*) of a long, straight alley (Anton starts from *A*, Ben starts from *B*). They each run at a constant speed and meet 800m away from the nearest starting point. They then continue running, and, after reaching the other end of the alley, they run back immediately and meet each other 400 m away from the other starting point. How long is the alley (see Figure 10.2)?

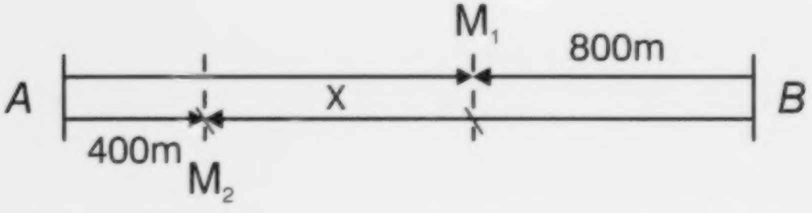


Figure 10.2

Again, we want to present two solutions that vary from each other in a similar way as in the above mentioned example. Maybe it should be emphasized that a sketch of the situation is very important for both solutions. A great part of the solution is already here in a suitable depiction.

Solution 1. The described situation may pose quite a dilemma for most students, as only very few things seem to be given and it may be not easy to find a "statement."

Let us suppose that Ben, who is starting from B, is the slower one. Then the first meeting point M_1 must be nearer to B than to A; $M_1B = 800$ m and $M_2A = 400$ m (Figure 10.2). It seems reasonable to let x be the length of the unknown piece, $M_1 M_2$ of $\overline{AB}$, the more so since the length of $M_2 M_1$ is required. However, neither of the two runners' speeds is known, nor are any times! As both runners started at the same time, it takes them the same time to get to the two meeting points, their
③ running times from the start to M_1 must be equal, and likewise from M_1 to M_2! Let v_1 be Anton's average speed and v_2 Ben's; then we get the following equations:

$$\frac{400 + x}{v_1} = \frac{800}{v_2}$$

$$\frac{800 + 800 + x}{v_1} = \frac{x + 400 + 400}{v_2}$$

But there are three variables in these equations! It is not possible to get three values from two equations, is it? Did we not think of all the information? Is this problem really solvable?

Indeed, in general, it is not possible to extract three variables (here x, v_1, and v_2) from two equations, and we did not forget anything. It is immediately clear that the runners' speeds cannot really be defined by the information given above: If both runners, for instance, were running twice as fast, then they both would be at the meeting points in half the time, but the meeting points would not change. So one can expect only a statement about the relationship of both speeds (but this is not really asked here.)

We surely cannot compute all three variables, but perhaps we could find at least one—how about the value of x? This is indeed possible, by dividing the first equation by the second equation (no occurring terms are zero); both speeds then disappear from the equation and we get an equation in x:

⑬

$$\frac{400 + x}{1{,}600 + x} = \frac{800}{800 + x}.$$

So we get the quadratic equation $320{,}000 + 1{,}200x + x^2 = 1{,}280{,}000 + 800x$, with the only positive solution being $x = 800$m. So the alley is 400m + 800m + 800m = 2,000m long. The relationship of the two speeds we get by inserting:

$$\frac{1{,}200}{v_1} = \frac{800}{v_2} \Rightarrow v_1 = \frac{3}{2} v_2$$

or, alternately, $v_1 : v_2 = 3 : 2$.

Solution 2. This solution does not need quadratic equations or, indeed, any equations at all, so it could be seen as particularly "elegant" in that way. Arithmetically, it contains, again, almost only elementary mathematics, but one has to see the situation in a suitable way:

"At their first meeting both sprinters together ran the whole length of the way exactly once." Have we already gained something by this? Not completely, but we've obtained a basis for the next steps! What is the situation at the meeting point M_2? At M_2, they have been running exactly three times the length of the alley (see Figure 10.2)! Now, as both sprinters run at a constant speed, not only has the total sprinted distance been tripled, but so have the individual distances run by each of the sprinters. Therefore, Anton and Ben are running from the start to M_2 three times the distance just to M_1. That means that Ben has to run a distance, from the start to M_2, of 3×800 m = 2400 m. Of that distance, he was running already 400 m back toward his starting point. Therefore, the length of the alley has to be 2400 m − 400 m= 2000 m! ③

2.3. Problems That Do Not Rely on Equations—Problems of Order

Although "text problems" and translating colloquial texts into the language of mathematics refer, respectively, in most cases to setting up and solving equations, not all described situations have to lead to equations. The latter kind of problems are extremely suitable for classes, as they stimulate the activity of the students and show the great variety of ways of thinking mathematically. Two connected examples follow.

Example 2.6

> There are six patients (*A*, *B*, *C*, *D*, *E*, and *F*) sitting in the waiting room of a dentist. The dentist knows (on the basis of previous treatments) what needs to be done for these particular patients and can therefore guess how long it will take him to treat each one. The treatment of *A* will last about 15 minutes; for *B* he estimates 30 minutes; for *C* and *D* approximately 10 minutes each; for *E* about 20 minutes; and for *F* only 5 minutes. The dentist wants to keep the total waiting time of all patients as low as possible and is thinking of an order in which to see the patients. In this sense, is there an optimal order? GEN

If he calls them, for instance, in the order *A*, *B*, *C*, *D*, *E*, *F*, so *A* does not have to wait at all, *B* would have to wait 15 minutes, *C* 45, *D* 55, *E* 65, and *F* 85. All together, this makes 265 min = 4 h 25 min. Are there more favorable orders? Is there, perhaps, even a most favorable one? If so, what is it? Is there a general rule for such problems? ⑭

Students could (should) try some orders, but it will hardly be possible to register all 6! = 720 possibilities in a systematic way.

How do the single waiting times and the total waiting time come about? Already in the example above it can easily be seen: If a patient Y has his turn after patient X, then the waiting time of Y obviously consists of the sum of the waiting time of X and the treatment time of X! The patients $P_1, P_2, \ldots, P_n$ have the times for their treatments $t_1, t_2, \ldots, t_n$. If they take their turns in that order, the following waiting times emerge:

⑬

P_1: 0
P_2: t_1
P_3: $t_1 + t_2$
. . .
P_{n-1}: $t_1 + t_2 + \ldots + t_{n-2}$
P_n: $t_1 + t_2 + \ldots + t_{n-2} + t_{n-1}$.

We add to get the total waiting time T:

$$T = (n-1)t_1 + (n-2)t_2 + \ldots + 2t_{n-2} + t_{n-1}.$$

The time for the treatment of the patient who is being treated first (t_1) is multiplied by the biggest factor $(n-1)$, thus giving it the "greatest weight" for the total waiting time, as this particular time has to be waited by all following patients. Now it can immediately be seen that if T shall become as small as possible, then the biggest factor $(n-1)$ has to be multiplied by the smallest possible value of t, $(n-2)$ has to be multiplied by the second smallest, and so on.

Result. The total waiting T time is the least if each time the patient with the shortest treatment time—among all still-waiting patients—has his turn first. The (one) optimal order for the initially asked problem is therefore $F \to C \to D \to A \to E \to B$, with a total waiting time of $T = 5 \times 5 + 4 \times 10 + 3 \times 10 + 2 \times 15 + 20 = 145$ min = 2 h 25 min.

So there is a waiting time saving of 2 hours compared with the alphabetical order—a considerable difference!

Remark. If there are patients with the same time for treatment, there is, of course, more than one optimal order. In our problem, the order $F \to D \to C \to A \to E \to B$ would also be an optimal one, as C and D have the same treatment time. We see that there may be more than one optimal solution to a problem, so, often, one cannot speak of *the*, but of *one*, solution.

Remark. Especially for young students, it is very important to learn that problems can have more than one solution. On the other hand, it is difficult to create such problems in an interesting way before equations of second or higher degree are taught. But certainly, this is too late!

An analogous problem, where one probably would have acted correctly intuitively, is the following:

As a winner of a lottery one can take bills from three piles of bills in the denominations \$100, \$50, and \$10. From any pile you choose first 10 bills, then from another pile 5 bills, and finally 1 bill from the third pile. To get as much money as possible, everybody would probably (without much thought) take the 10 bills from the \$100 pile, the 5 bills from the \$50 pile, and 1 bill from the \$10 pile. The sum of the winnings is maximal if the biggest factor (10) is multiplied by the largest value (\$100). The sum of the winnings corresponds here with the total waiting time (but as big as possible instead of as small as possible), and the values \$100, \$50, and \$10 correspond to the waiting times. The order of the piles (with the fixed order of the taking of the bills 10 → 5 → 1) corresponds to the order of the patients.

Now a second problem of that kind (generalization!).

Example 2.7

A senator has to welcome five special-interest groups for brief talks (separately, one after the other). The members of the groups are already in the waiting room, and every group already has its fixed time for the consultation (compare with Table 10.1). In what order should the senator call in the groups in order to keep the total waiting time of all the individual members as low as possible?

GEN

The situation is again similar to the one at the dentist's office (but now with groups rather than individuals).

Table 10.1

Group	*Number of Members*	*Consulting Time*
1	4	20 min.
2	8	10 min.
3	5	30 min.
4	10	15 min.
5	6	25 min.

⑥ ⑭

Solution 1. If group 1 (consisting of 4 members) is, for instance, in consultation with the senator for 20 minutes, it can be considered the same as if each of the these 4 persons were there for 5 minutes (20/4 = 5). So it is possible to split the groups in single persons and the group times into single times (times per person). The above example gives a solution to the problem of an optimal order of single persons: A person with the shortest possible consultation time has to take his or her turn first. So the times per person determines the single order and in the sum, therefore, also determines the order of the groups, because in an optimal order of the individuals, the members of a group would certainly have their turn for the consultation one after another according to the above principal and could be seen in this respect as a group again. In other words, to call all persons with the currently shortest consulting time into the room is the same as calling in the whole group in its entirety first!

The consulting time per person in the single groups is indicated in Table 10.2:

Table 10.2

Group	1	2	3	4	5
Time per Person (min.)	5	1¼	6	1½	4⅙

Therefore, the optimal order for the groups, according to the increasing times per person, is $G_2 \rightarrow G_4 \rightarrow G_5 \rightarrow G_1 \rightarrow G_3$.

Solution 2. The following would be a slightly different way of dealing with the
③ problem (for instance, in the case where the example above was not used). Let us reduce the problem first to the two groups G_1 and G_2. Is it better to take G_1 first or G_2?

⑰ If G_1 has its turn first, then the 8 persons of G_2 have to wait 20 minutes each (i.e., the consulting time of G_1), with a total waiting time of 160 minutes. If G_2 were to go in first, the 4 persons of G_1 have to wait 10 minutes each, with a total waiting time of 40 minutes, which is considerably less.

⑬ Let us now consider the general problem again: n different groups $G_1, \ldots, G_n$ with $g_1, \ldots, g_n$ members and the consulting times $t_1, \ldots, t_n$ are sitting in the waiting room. Let us assume they take their turn in the order $G_1, \ldots, G_k, G_{k+1}, \ldots, G_n$ to see the senator, and this results in a total waiting time T_1. What happens if any two consecutive groups are interchanged? If one interchanges, for instance, the two groups G_k and G_{k+1}, we get the order $G_1, \ldots, G_{k+1}, G_k, \ldots, G_n$ and a new total waiting time T_2. There is obviously no change in the waiting times of any of the groups not involved in this particular change ($G_1, \ldots, G_{k-1}$ and $G_{k+2}, \ldots, G_n$); nor does the combined waiting time that G_k and G_{k+1} wait together for $G_1, \ldots, G_{k-1}$ change. Let the sum of these (in every case) constant waiting times be S. Then T_1 consists of S and the waiting time of G_{k+1} for G_k: $T_1 = S + t_k g_{k+1}$

T_2 is derived analogously if G_{k+1} has its turn before G_k: $T_2 = S + t_{k+1} g_k$.

Under what conditions is T_1 smaller than T_2? Obviously, when $t_k g_{k+1} < t_{k+1} g_k$, which is equivalent to $\frac{t_k}{g_k} < \frac{t_{k+1}}{g_{k+1}}$.

So if t_k/g_k (the consulting time per member of the group G_k) is smaller than t_{k+1}/g_{k+1}, it is more favorable to ask the group G_k to the consultation first.

Also from this we can deduce in general: If it should occur in a certain order that the inequality $\frac{t_k}{g_k} > \frac{t_{k+1}}{g_{k+1}}$ holds true for any k, it is possible to reduce the total waiting time by exchanging the two groups.

Result. The order is optimal if and only if $\frac{t_1}{g_1} \leq \frac{t_2}{g_2} \leq \ldots \leq \frac{t_n}{g_n}$. So the result here as well is that the consulting time per person is decisive in finding the best order.

Especially for such problems, it is extraordinarily important to give the students enough time for presumptions and trials and to talk about all suggestions and arguments. When necessary, the students may be given little hints—the teacher should not disclose too much and should encourage students to work independently. The teacher, of course, has to lead or channel all problem-solving processes to a certain degree, since the students will probably often choose wrong approaches or will just not know what to do next. But this leading should not take the form of an especially elegant solution that appears from nowhere.

3. Problems in the Context of Special Mathematical Theories

This section will deal with some problems that illustrate the essence and the importance of certain mathematical themes and that contain easily recognizable "typical features" of these theories. Such (or similar) examples may sometimes even serve as an entry into certain mathematical topics if they are suitable for discovering the essentials of the new topic and if they motivate students to deal thoroughly with the particular issue. Such entry problems must not be too sophisticated and should show clearly that an increase in knowledge (e.g., precision, theoretical investigation, systematization, etc.) is necessary or at least possible.

3.1. A Problem Representing Different Theories

Example 3.1

GAME

"The Tower of Hanoi" (e.g., Haussmann, 1986; Stowasser & Mohry, 1977): A tower of *n* (e.g., *n* = 3) discs is built on one of three rods, with the discs getting smaller from bottom to top (see Figure 10.3).

The task is now to move this "*n*-tower" from one rod to another while obeying two rules:

a. Only one disc is allowed to be moved from one rod to another each time.

b. A larger disc may never rest on a smaller one.

What is the minimal number *H(n)* of single steps (moving of discs) for a tower of *n* discs?

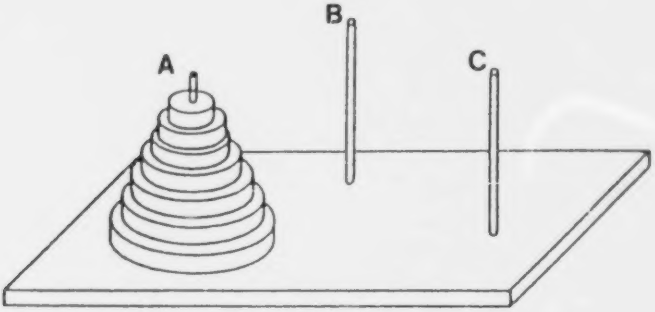


Figure 10.3 The tower of Hanoi

It is certainly a great advantage to have a few real trials with "model" towers of objects that get smaller (e.g., books, pieces of paper). By trial and error one can immediately make the following table for small *n:*

Table 10.3

②

n	1	2	3	4
H(*n*)	1	3	7	(15)

Many students will already begin having problems with *n* = 4, but for students with even a little interest, *H(4)* = 15 will not be too hard to work out. But now the question seems to be quite obvious, whether there is a "system," and whether it is possible to get *H(4), H(5), H(6),* . . . by a thought experiment (e.g., from the known values of *H(1), H(2),* and *H(3),* respectively). Let us go back one more step: Is it already possible to get *H(3)* = 7 by consideration, without really experimenting? ③ Put a different way: How can I move a 3-tower if I know how to move a 2-tower?

One can generally determine an answer easily from the concrete experience (18) in experimenting: To move a 3-tower from rod *A* to rod *C*, one first has to move the upper 2-tower to rod *B*, then the lowest (largest) disc from *A* to *C*, and finally the 2-tower "stored" in *B* to *C* (and that by employing the same scheme used to move it before from *A* to *B;* see Figure 10.3). As it is known how a 2-tower has to be moved and how many steps it takes—*H(2)* = 3—it follows that *H(3)* = *H(2)* + 1 + *H(2)*; 7 = 3 + 1 + 3, as the experiment confirmed already.

Now the structure of the problem becomes clearly visible: The same principle must now be applicable for $H(4)$: $H(4) = H(3) + 1 + H(3) = 2 \times H(3) + 1$. Analogously, to move a 4-tower, one has to move a 3-tower first, then the lowest disc, and finally again the 3-tower. So now it is not especially hard to discover: $H(n) = 2H(n-1) + 1$.

The value $H(n)$ is easy to determine if $H(n-1)$ is known, this in turn if $H(n-2)$ is known and so on, until one gets to a known $H(k)$. By this, one can ⑨ determine step by step: $H(k+1) = 2H(k) + 1$, $H(k+2) = 2H(k+1) + 1, \ldots, H(n) = 2H(n-1) + 1$.

In this manner, the above table can be conveniently continued: The last number with which it seems sensible to conduct the experiment in reality is $n = 6$, because starting at $n = 7$ the number of necessary single steps is already very high.

Table 10.4

n	1	2	3	4	5	6	7	8	9	10	...
H (n)	1	3	7	15	31	63	127	255	511	1,023	...

One can see that this example very naturally leads to the principle of recursion or iteration. So it can be a gateway to the accompanying theory if the student has not heard of it. Another theory demonstrated here would be that of ALG the difference equations (such as linear difference equations with constant coefficients), which are increasingly important for the description of dynamic systems, especially those of first order ($a_n = ka_{n-1} + d$) and perhaps those of second order ($a_n = k_1a_{n-1} + k_2a_{n-2} + d$).

The principle of proof by mathematical induction can also be illustrated and even motivated by this problem. Generally, it seems that proofs by induction ALG especially are often introduced without any natural motivation, without "discovery" ⑱ and through a relatively dry lecture that doesn't help students really understand the necessity or the principle itself!

For big n (e.g., $n = 50$), of course, this step-by-step calculation of the values of $H(n)$ is relatively clumsy (at least without a programmable calculator or computer). So the question arises whether it is possible to express $H(n)$ directly by a formula without having to start at $H(2)$ or $H(3)$ and coming forward only gradually. For this we should look again at Table 10.4. The powers of 2 (2, 4, 8, 16, 32, 64, 128, . . .) are very significant and impressive numbers, so one will soon notice that the values of $H(n)$ are always 1 smaller than 2^n. Therefore, the assumption : $H(n) = 2^n - 1$ seems justified. According to Table 10.4, it is true at least for $n \leq 10$.

Now, what does the case $n = 11$ look like? We know $H(10) = 2^{10} - 1$, and we further know that $H(11) = 2(2^{10} - 1) + 1 = 2 \times 2^{10} - 2 + 1 = 2^{11} - 1$.

So our assumption is true as well for $n = 11$. We proved the validity of the formula for $n = 11$ by the validity for $n = 10$. For $n = 12$ we therefore get $H(12) = 2H(11) + 1 = 2(2^{11} - 1) + 1 = 2 \times 2^{11} - 2 + 1 = 2^{12} - 1$.

One can already see that the validity of the formula is "inherited" from one number to another (one could also use further concrete examples). It is already perceptible, from these concrete examples, that this hereditary character does not depend on the particular choice of n (e.g., $n = 11$ or $n = 12$). We could repeat the above step to prove thereby the validity of the formula for any fixed n. But it is much more convenient to show this hereditary character in general (i.e., detached from concrete numbers): We want to show the heredity of the formula for all $n \in N$. Let us assume the validity of the formula $H(n) = 2^n - 1$ for any n. It then follows that $H(n+1) = 2H(n) + 1 = 2(2^n - 1) + 1 = 2^{n+1} - 1$, and this is the validity of the formula for the next number $n + 1$.

But if a formula is true for $n = 1$, and if the validity for any natural number ensues from the validity for the previous natural number, then the formula must obviously be true for all natural numbers:

$$H(1) \rightarrow H(2) \rightarrow H(3) \rightarrow H(4) \rightarrow H(5) \rightarrow \ldots$$

We believe with this example, students could discover the theories of recursion and complete induction almost by themselves in a genetic way, and could therefore perhaps understand them better (see Stowasser & Mohry, 1977, 7ff).

③ Other possibilities of proving the validity of $H(n) = 2^n - 1$ would be, on the one hand, through the sum formulas of geometric series (another theory) or, on the other hand, by the following consideration:

We know the recursion equation $H(n+1) = 2H(n) + 1$ with $H(1) = 1$ is true. From that we get, step by step:

$$\begin{aligned}
&H(2) = 2 \times 1 + 1 = 2 + 1 = 2 + H(1)\\
&H(3) = 2(2 + 1) + 1 = 2^2 + 2 + 1 = 2^2 + H(2)\\
&H(4) = 2(2^2 + 2 + 1) + 1 = 2^3 + 2^2 + 2 + 1 = 2^3 + H(3)\\
&\ldots\\
&H(n+1) = \ldots = 2^n + 2^{n-1} + \ldots + 2 + 1 = 2^n + H(n).
\end{aligned}$$

By equating both terms for $H(n+1)$, one gets $2H(n) + 1 = 2^n + H(n)$, from which immediately follows $H(n) = 2^n - 1$.

3.2. *Simple Tasks to Be Solved by Graphic Representation*

A special device for the solution of some problems is graphic representations. Drawing sketches often initiates and supports the process of problem solving and takes it in the right direction (see examples 2.2 and 2.5). An example of a graphic illustration making an important contribution to the solution comes in network technique.

Network technique is a relatively young method (developed in the late 1950s) for better planning, controling, coordinating, and supervising of bigger enterprises or projects. One of the most essential aims of this technique is to determine the earliest possible completion date of a project, which consists of several "partial tasks."

Mathematics can be seen in a certain way as a language, and its special structure can contribute decisively to the solution of many problems. Its means are, among others, variables, equations, and graphics (illustrations). ⑥ Graphic presentations of all kinds especially serve to describe situations as they can also be described by variables, equations, formulas, and the like; but graphic presentations can often be realized better, more easily, and more quickly than texts or representations with variables. "A picture *is* worth a thousand words."

The network technique is an excellent opportunity to make these graphic aspects clear. Here, not only are graphic representations worked out, interpreted, and described, but there is an essential additional factor: One can also see this technique as an "intelligent" graphic representation in the sense that logical connections, formalized relations, quantitative connections, and the like are illustrated in such a way that the representation already gives a huge part of the solution. Here is a very simple example that can be especially suitable as an entry-level problem.

Example 3.2

A small construction project is divided into certain stages $P_1, P_2, \ldots, P_6$ that are achieved by the completion of single jobs. P_1 will be the beginning stage and P_6 the end stage. ALG The notation $\overline{P_2\,P_5} = 9$ means that P_2 is a precondition for P_5, and the work that has to be done "between" P_2 and P_5 takes at least 9 days (time units). The complete information on the whole project is given by:

$$\overline{P_1P_2} = 4;\ \overline{P_2P_3} = 6;\ \overline{P_2P_5} = 9;\ \overline{P_2P_4} = 8;\ \overline{P_3P_5} = 7;\ \overline{P_5P_6} = 3;\ \overline{P_4P_6} = 6.$$

The lengths of the single working processes $P_i \to P_j$ is based on experience. What is the minimal length of time for the entire project?

Solution. The given information about the single $\overline{P_iP_j}$ contains information both about which stages have to be reached before others and about how long each single working process lasts. The graphic translation of the above information leads to a netlike representation (Figure 10.4) called a "network," in which the single working process is marked as an arrow between the single stages. Such an representation gives a good survey of the situation and allows, in simple cases, a solution by trial and error. (See Chapter 15.) ⑪

Since, for completion, all processes have to be done, one has to look for the longest path from P_1 to P_6 to get the shortest total time. The other processes can then easily be "fit in." This may present a slight paradox for some students. They intuitively might have looked for the shortest path, but it is not difficult to see that the jobs that last longer could not be completed then. By trying all possible paths from P_1 (start) to P_6 (end)—adding all the working times along the paths—in simple "networks" such as the one in Figure 10.4, one will quickly see that here the "critical path" is $P_1 \rightarrow P_2 \rightarrow P_3 \rightarrow P_5 \rightarrow P_6$ ("critical" only because the jobs along this path obviously must not be delayed to prevent an extension of the shortest total time). The processes $P_2 \rightarrow P_5$ and $P_2 \rightarrow P_4 \rightarrow P_6$ can be fit in parallel fashion, as they are shorter. The shortest time for the realization of the project (e.g., construction time) is therefore $4 + 6 + 7 + 3 = 20$ time units.

⑥

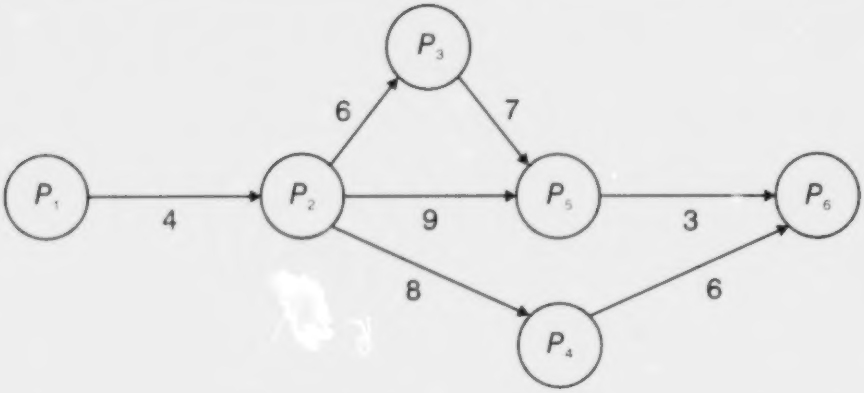


Figure 10.4 Network of a small project

Remark. If students have never been confronted with this kind ot task, even simple ones such as the above can pose a problem in the sense that students have to look for a (for them) new kind of solution.

Also, in slightly more complicated "networks" (e.g., Figure 10.5), it is possible to find one solution by trying different paths. As this example shows, there may be several solutions (i.e., the critical or longest paths). In the network of Figure 10.5, the two paths $P_0 \rightarrow P_1 \rightarrow P_5 \rightarrow P_7$ and $P_0 \rightarrow P_3 \rightarrow P_4 \rightarrow P_7$ are both critical ways with a total length of 15 time units each (see Reichel et al., 1989-1992, p. 171).

For the solution of very complex building projects, of course, this method of trial and error will not be sufficient anymore. Some algorithms have already been developed to determine the critical (longest) paths with the help of highly efficient computers. These algorithms are not at all so complicated that they could not be dealt with in school. There is really a theory behind such tasks, and even students can gain a first insight into it.

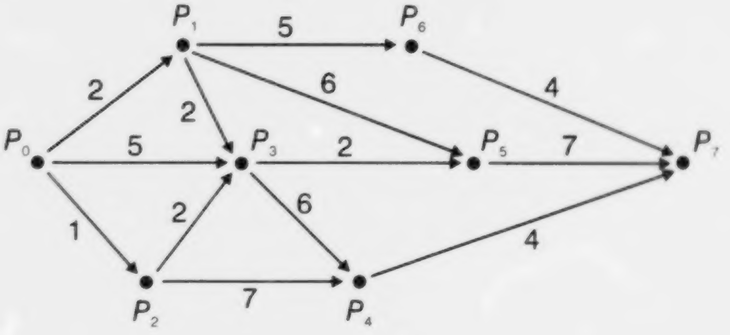


Figure 10.5 Example of a network

3.3. *Problems Concerning Rational and Irrational Numbers*

We do not know if all students develop the right idea of the importance and the nature of irrationality by the current examples in school (e.g., proof of the irrationality of $\sqrt{2}$). What can it mean or what effect does it have if a number cannot be expressed as a fraction (i.e., the relation of two whole numbers)? The following problems, along with the solutions, will show some surprising results.

Example 3.3

In a system of coordinates, a point with integer coordinates is called a *lattice point* (the point [4, 3], for instance, would be a lattice point, and the point [0.5, 2] would not).

Is there a straight line through the origin of a two-dimensional system of coordinates that does not contain any lattice point (Figure 10.6)? For reasons of symmetry, one can confine attention here to positive coordinates.

ALG

⑥

Of course there are many straight lines that do not contain a lattice point in the shown sector of Figure 10.6, but whether they do not meet a lattice point "anywhere" is surely an exciting question!

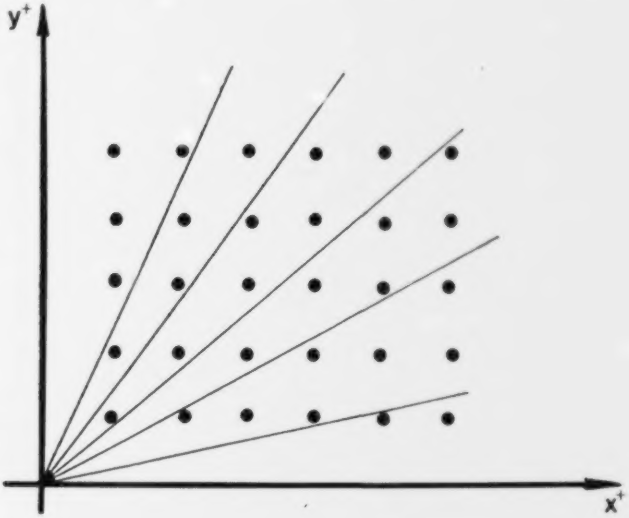


Figure 10.6 Lines and lattice points

Another formulation could be: Is it possible to fire from the origin an idealized "arrow" (with diameter zero) through a "forest" with "trees" at the lattice points (also with diameter zero) without ever hitting a "tree" (the "arrow" may eternally maintain its original direction)?

Solution. If a straight line contains a lattice point with the coordinates (m, n) m, $n \in N$, it has the slope n/m. This means that if it is not possible to express the slope of a straight line as a quotient of two natural numbers, then it cannot contain any lattice points. The inverse is, of course, true as well: If a straight line has the slope
⑬ n/m, it will contain a lattice point, such as (m, n). So only the straight lines with irrational slopes do not contain lattice points, and for this there are many more possibilities than for straight lines with rational slopes. Although there is an infinitude of straight lines with both rational and with irrational slopes, one can formulate: "The majority of straight lines through the origin do not contain lattice points" (noncountable infinitude in contrast to countable infinitude of straight lines with lattice points).

To continue, it could be shown: If a straight line through the origin contains one additional lattice point (other than [0, 0]) it must contain an infinite number of them! What is the distance between two neighboring lattice points?

Solution. If (a, b) is a lattice point, then $(2a, 2b)$, $(3a, 3b)$, . . . are, of course, lattice points, too. These are neighboring points if a and b are relatively prime. In this case we get, for the distance, $\sqrt{a^2 + b^2}$.

Example 3.4

> Take any circle and mark a point P_0. Then translate (rotate) P_0 by the angle α to obtain the point P_1. Repeat the process to obtain P_2, P_3, P_4, . . . (see Figure 10.7). One of the points P_n may be equal to point P_0 again. In that case, our sequence of points is finite. On the other hand, the sequence obtained by the process described can be infinite, too. Then it consists of pairwise different points, of course. Now the question arises: For which angles α do we get finite sequences, and for which α infinite ones? In other words, for which α is there an index n such that $P_n = P_0$?

ALG

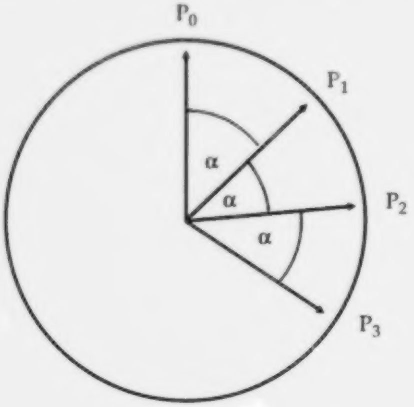


Figure 10.7 Finite or infinite sequence

Solution. If $P_n = P_0$, for some n, some multiple of α is a multiple of 360°, too; in other words, there are natural numbers k and l so that the following is valid:

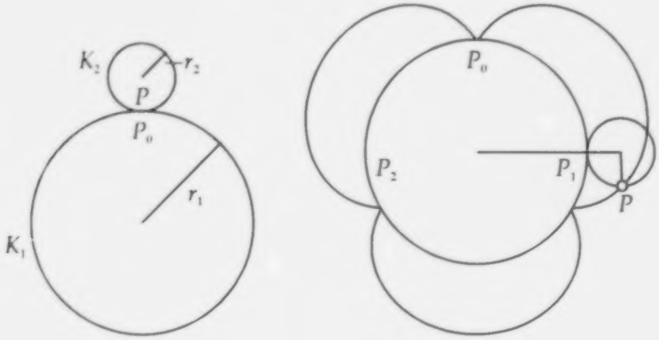
$$360 \times k = \alpha \times l \iff \alpha = \frac{360k}{l}.$$ ⑬

This means α has to be rational. Conversely, if α (measured in degrees) is rational, the sequence $P_0, P_1, P_2, P_3, \ldots$ is finite. Since, if $\alpha = a/b$ $(a, b \in N)$, we have $(360° \times b) \times \alpha = a \times 360°$. In other words, $P_0 = P_{12}$ for $n = 360° \times b$. In conclusion, $P_0, P_1, P_2, P_3, \ldots$ is infinite if and only if α (measured in degrees) is irrational.

Example 3.5

GEOM

A circle K_1 with the radius r_1 touches another circle K_2 with the radius r_2, and the points that touch are called $P_0 \in K_1$ and $P \in K_2$, respectively (see Figure 10.8a). By rolling along the circumference of K_1, the point P describes an arched curve called an *epicycloid* (see Figure 10.8b), where the "next" points that touch (when P meets again exactly with K_1) are called $P_1, P_2, \ldots$ Under which conditions of r_1 and r_2 will the curve of the point ever close? (When will the point P ever return to the starting point P_0? When is there some $r \in N$ with $P_n = P_0$?)



a) b)

Figure 10.8 Rolling of a circle—"epicycloids"

Solution. Here one could rush to the conclusion: The curve will be closed only if r_1 and r_2 have rational values, otherwise not. The curve will obviously be closed if and only if both circumferences (u_1 and u_2) have a common multiple—that is, if there are natural numbers k and l with $(2r_1\pi) \times k = (2r_2\pi) \times l$. This is equivalent to ⑬ $r_1 \times k = r_2 \times l$ or, alternately, $r_1/r_2 = l/k$. Therefore, it is not the irrationality of r_1 or r_2 individually that is decisive here, but that of u_1/u_2 or r_1/r_2. If, for instance, $r_1 = 3\pi$ and $r_2 = \pi$, then the curve obviously closes exactly after one "round" of K_1 or, equivalently, after three resulting "arches" (e.g., see Figure 10.8b), though both radii have irrational values.

Result. If the fraction r_1/r_2 is irrational, then the curve of P will never close. If r_1/r_2 is rational, so that $r_1/r_2 = m/n$ (m and n relatively prime), then the curve will close after exactly n rounds of K_1, or equivalently, after m arches.

To better appreciate the full flavors of rationality, we shall consider a few more problems.

Example 3.6

Are there positive numbers $a, b \in Q$, but $a, b \notin N$, so that $a^b = c \in N$? ALG

Most students will know the following proposition: The square root of a natural number is either again a natural number or an irrational number. The same holds true for the k^{th} root instead of the square root (to be proved analogously).

Solution. There are no such numbers a and b. ⑬

Indirect proof: Let $a = m/n$ ($a \notin N$) and $b = k/l$ ($b \notin N$) then the above condition states

$$\left(\frac{m}{n}\right)^{\left(\frac{k}{l}\right)} = c \ (c \in N).$$

This yields:

$$\left(\frac{m}{n}\right)^k = c^l \text{ or, equivalently, } \frac{m}{n} = \sqrt[k]{c^l}$$

As $c^l \in N$, then, according to the above proposition, $\sqrt[k]{c^l}$ is either a natural number itself (contradiction to $m/n = a \notin N$) or an irrational number (contradiction to $m/n = \sqrt[k]{c^l}$). That is why such numbers cannot exist.

Example 3.7

Are there $a, b \in R$ but $a, b \notin Q$, so that $a^b = c \in Q$ (cf. Heinze, 1993, 147f.)? ALG

Solution. $\sqrt{2}$ is surely an irrational number. Consider, for instance, $\sqrt{2}^{\sqrt{2}}$—one of the simplest expressions of the demanded form—and let us ask if this value is ⑬ rational or not. If this value is rational, then we have already found one example and have to answer the question posed with Yes; if not, we have to continue working. So let us assume that $\sqrt{2}^{\sqrt{2}}$ is not rational and continue trying. We form, for instance,

$$(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{2} = 2$$

and see that the result, 2, is not only rational but also a natural number, though the exponent ($\sqrt{2}$) is irrational and (according to the assumption) the base ($\sqrt{2}^{\sqrt{2}}$) is irrational, too.

Result. If $\sqrt{2}^{\sqrt{2}}$ is rational, then we have to say "yes" to the question posed; if it is irrational, then with $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$ we have found an example of the demanded form and have to answer this question as well with Yes. Even without knowing whether $\sqrt{2}^{\sqrt{2}}$ is rational or irrational, we have found an example in every case.

Analogous examples would be $(\sqrt{5}^{\sqrt{2}})^{\sqrt{2}}$ or $(\sqrt[3]{2}^{\sqrt{3}})^{\sqrt{3}}$.

3.4. *Problems Concerning Complex Numbers*

The following problems have the theory of complex numbers as a background and contain some surprising results. Especially surprising moments in class often promote motivation as well as fascination, excitement, and interest. But it should be emphasized that we are dealing here with problems that will find successful application in more advanced classes. The first problem is a continuation of example 3.7.

Example 3.8

ALG Are there $a, b \in C$, but $a, b \notin R$, with $a^b = c \in R$?

⑬ Students generally learn Euler's formula $e^{i\phi} = \cos\phi + i\sin\phi$ without an exact foundation by expansion into series. But the exact derivation and the knowledge of all problems around the exact definition of the complex exponential function are not really necessary for the following considerations.

It is quite simple to insert some values into Euler's formula and just wait with excitement for the result. For $\phi = \pi$ one can already get the first interesting result: $e^{i\pi} = \cos\pi + i\sin\pi = -1$.

The numbers e and π are irrational (even transcendental), $i \times \pi$ is a complex number, and $e^{i\pi}$ is nevertheless an integer—really surprising! For $\phi = 2\pi$ one gets $e^{2\pi i} = 1$—a natural number.

Now let us take $\phi = \pi/2$ and substitute it; we get $e^{i\frac{\pi}{2}} = i$—also an interesting identity. The simplest power where base and exponent are not real numbers is i^i. We get

$$i^i = \left(e^{i\frac{\pi}{2}}\right)^i = e^{i^2\frac{\pi}{2}} = e^{-\frac{\pi}{2}} = \frac{1}{\sqrt{e^{\pi}}},$$

a real number, which is expressed again by the numbers e and π, which are extremely important for mathematics in general.

With help of the Euler's formula we can also get an insight to the periodicity of the complex exponential function and can find as well solutions to problems like the following.

Example 3.9

Does the equation $1^x = 3$ have a solution in *C?* ALG

As seen already, Euler's formula yields, for $\phi = 2\pi$, the length of the identity $e^{2\pi i} = 1$. Since the functions sin and cos are periodical with a period 2π, the formula also holds true for $\phi = 4\pi, 6\pi, \ldots, 2k\pi, \ldots$: ⑬

$$e^{2k\pi i} = \cos 2k\pi + i \sin 2k\pi = 1 \ (k \in Z)$$

We see the equation $e^x = 1$ has not only the solution $x = 0$ but even an infinitude of solutions $x = 2k\pi i$ $(k \in Z)$. Now to the equation of the posed problem $1^x = 3$, which can be described—as we have just seen—in the following way, too:

$$(e^{2k\pi i})^x = e^{ln3} \Leftrightarrow e^{2k\pi i\,x} = e^{ln3},$$

from which immediately follows:

$$x = \frac{ln3}{2k\pi i} = -\frac{ln3}{2k\pi} \times i \ (k \in Z).$$

The given equation does not have a solution in *R*, but in *C* it even has infinite solutions, so this is an especially interesting proof of the "mathematical abundance" of complex numbers. Of course, one cannot carry on complex analysis in school, but a little insight into the possibilities of mathematical extension by complex numbers seems to us very interesting for students and beneficial to their motivation, especially in advanced courses.

Analytic geometry also offers some surprises and good opportunities to incorporate complex numbers into certain considerations.

Example 3.10

ALG

Given any circle and a point P outside of it, we easily can find the equation of the tangent through P. But what if P lies inside the circle. Is there still a tangent? Formally, the point P can be inserted into the tangent-equation. But is there any meaning in doing so?

⑬ It surely needs no explanation that this is not possible in the range of real numbers. But in the realm of complex numbers, even this is possible! We want to try, for instance, the point with coordinates (0, 0), the center-point of the circle $x^2 + y^2 = r^2$, to put formally/arithmetically the tangents on that circle. From analytical geometry we know that the line $y = kx + d$ will touch the circle $x^2 + y^2 = r^2$ if and only if $r^2(1 + k^2) = d^2$.

As the tangent shall be put on the circle from the point (0, 0), $d = 0$, and because of $r \neq 0$ it follows: $1 + k^2 = 0$, so $k = \pm i$.

Therefore, the equation of the tangents is $y = \pm ix$. As expected, these are imaginary lines, purely complex ones. These straight lines have some other surprising properties as well.

Which straight lines are orthogonal to $y = ix$ and $y = -ix$, respectively?

It is well known that two straight lines are orthogonal if for their slopes k_1 and k_2, the following is valid: $k_1 k_2 = -1$. For the straight line $y = ix$, $k_1 = i$; what is wanted now is a k_2 with $ik_2 = -1$ (i.e., $k_2 = i$). Therefore, the straight line that is orthogonal to $y = ix$ is $y = ix$ itself (!)—and this is not a misprint; in the range of complex numbers, there are straight lines that are orthogonal to themselves (analogously: $y = -ix$).

Take any pair of points P_1 and P_2 of the line $y = ix$ (or $y = -ix$, respectively). What distance will they have?

$P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ will be two points on the line $y = ix$. The length (Euclidean distance) of segment $\overline{P_1P_2}$ always equals zero:

$$(P_1P_2)^2 = (x_2 - x_1)^2 + (ix_2 - ix_1)^2 = (x_2 - x_1)^2 + i^2(x_2 - x_1)^2 = 0. \text{ (Since } i^2 = -1\text{)}$$

For a successful treatment of these topics, at least two preconditions have to be fulfilled:

1. The performance of the class must be at a relatively high standard, because the processes of abstraction that have to be executed are not really elementary.
2. The teacher has to be excited by the surprising results and also has to be able to transmit this excitement (tension) to the "positive attitude of expectation" of the students.

3.5. *Problems Concerning Sums of Natural Numbers*

The following problems refer to the field of "arithmetical series," especially to the formula

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

This formula in particular is often practiced only by using stereotypical and recipelike tasks. The following problems are a little more multifarious.

Example 3.11

Calculate the sum of all natural numbers up to 300 that are divisible by neither 8 nor 6. ALG

This is a typical task that can easily be divided into subtasks. First one could calculate the sum of all natural numbers from 1 to 300:

$$1 + 2 + 3 + \ldots + 299 + 300 = \frac{300 \times 301}{2} = 150 \times 301 = 45{,}150.$$ ⑬

Now the sum of those numbers that are divisible by 8 or 6 is to be subtracted. The sum of all numbers (≤ 300) that are divisible by 8 is

$$8 + 16 + 24 + \ldots + 296 = 8(1 + 2 + 3 + \ldots + 37) = 8 \times \frac{37 \times 38}{2} = 5{,}624.$$

The sum of all numbers (≤ 300) that are divisible by 6 is:

$$6 + 12 + 18 + 24 + \ldots + 294 + 300 = 6 \times \frac{50 \times 51}{2} = 7{,}650.$$

If one now forms the sum 5,624 + 7,650 = 13,274, then one gets the sum of all numbers that are divisible either by 8 or by 6—or do we? No, there has to be a little correction: The number 24, for instance, was counted twice (see above), because it is divisible both by 8 and 6. We see that all numbers with this quality were counted twice, which means we have to subtract them once. Which numbers are these that are divisible by 6 and 8? It is not particularly difficult to see that they are the multiples of 24.

Therefore the sum

$$24 + 48 + 72 + \ldots + 240 + 264 + 288 = 24(1 + 2 + \ldots + 11 + 12) = 24 \times \frac{12 \times 13}{2} = 1{,}872$$

has to be subtracted from 13,274: 13,274 – 1,872 = 11,402. Now this number has to be subtracted from 45,150, and as the result we get 33,748.

Example 3.12

ALG Beginning with 1, the natural numbers are added until one gets a three-digit number with three equal digits. How many numbers have to be added (cf. Baron & Windischbacher, 1990, pp. 14, 89f)?

③ *Solution.* One solution would be to simply add up step by step until one gets the demanded form. But we want to choose another method. We have to add n numbers, and this sum

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

⑬ may not exceed 999. By inserting, or by solving a quadratic equation, one can see easily that n may not exceed 44. We look, therefore, for an $n \leq 44$, so that

$$\frac{n(n+1)}{2} = xxx \qquad (1 \leq x \leq 9)$$

Now, unfortunately, the expression "xxx" is not very usable; therefore, we are looking for a more suitable and simpler expression of "xxx," such as $xxx = x \times 111$, and therefore

$$\frac{n(n+1)}{2} = x \times 111.$$

But the number 111 can be written even more basically as a product of primes $37 \times 3 = 111$; then we get

$$n(n+1) = 2 \times x \times 37 \times 3 \qquad n \leq 44; \qquad 1 \leq x \leq 9.$$

As 37 is a prime, either n or $n + 1$ must be divisible by 37. So with the condition $n \leq 44$, only the possibilities $n = 37$ or $n = 36$ remain. For $n = 37$ we get $37 \times 38 = 2 \times x \times 37 \times 3$, but this is not possible, as 37×38 is not divisible by 3. Therefore, the only remaining possibility is that $n = 36$. From $36 \times 37 = 2 \times x \times 37 \times 3$, it follows immediately that $x = 6$. The sum of the numbers from 1 to 36 is really 666.

The method used here (factorization of 111 as a product of primes) may look like a "trick." But when teaching problem solving, we always use such tricks that—and this is the most important item—turn out to be mathematical methods as soon as we recognize that there are common features and structures of problems corresponding to several kinds of "tricks." Factorization into a product of primes is a significant example of that!

Example 3.13

Determine the sum of the digit-sums of all natural numbers from 1 to 999. ALG

Solution. What does it mean, to determine sums of digit-sums of certain numbers? ③
The digits are obviously counted (summed) as often as they occur in the single numbers; for example, the sum of the digit-sums of the numbers 123, 203, 125, and ⑬
52 is $(1 + 2 + 3) + (2 + 3) + (1 + 2 + 5) + (5 + 2) = 1 \times 2 + 2 \times 4 + 3 \times 2 + 5 \times 2 = 26$.

So we have to rephrase the question to ask how often the digits 1, 2, . . . , 9 occur in the numbers from 1 to 999! It is irrelevant how often the digit 0 occurs, as it does not contribute to the digit-sum. First of all it is clear (or could become clear after some concrete considerations) that the digits 1, 2, . . . , 9 occur equally often in the numbers from 1 to 999, as each number appears equally often in each place (hundreds, tens, and ones), but how often? Let us split the problem again. How often, for instance, does the digit 3 occur as a hundred-digit? Obviously 100 times (300-399). How often does the 3 appears as a ten-digit? In every "hundred section" 10 times (e.g., 130-139). Since there are 10 of these "hundred sections," the digit 3 occurs altogether $10 \times 10 = 100$ times as the ten-digit. As the one-digit, the digit 3 occurs as well 10 times in every "hundred section" (e.g., 103, 113, 123, . . . , 193), so altogether 100 times as well. We see that the digit 3 (as well as all the other digits, except 0, which occurs altogether 189 times) occurs 300 times altogether in the numbers from 1 to 999. The sum of the digit-sums is therefore

$$300(1 + 2 + \ldots + 9) = 300 \times \frac{9 \times 10}{2} = 13{,}500.$$

3.6. *An Interesting Problem Concerning the Divergence of the Harmonic Series*

The following example from Kranzer (1989, pp. 157, 223ff) should not be interpreted as a practice-oriented example. The wording here is only one aspect that should support the surprise from the result of this thought experiment and may contribute to motivation.

Example 3.14

ALG

A snail is sitting at the end *A* of a 1 km long and arbitrarily elastic rubber band. The snail has only one task in its infinitely long life: to creep at a speed of 1 cm/s toward the other end of the rubber band. After each second, the rubber band is stretched by 1 km by an invisible power (homogeneous stretching). Will the snail ever reach the other end *B* of the rubber band? If yes, how long will it take it?

⑬ *Solution.* Let us first consider how the length of the band changes in the course of the time. At each full second the band is homogeneously stretched by 1 kilometer. Let the time immediately before the completion of the n^{th} second be n^-, and the one immediately after the completion of the n^{th} second be n^+; between both the band has been stretched by 1 km. At the time n^-, the band is n km long, but at the time n^+, its length is $n + 1$ km, caused by the consistent homogeneous stretching of the whole band by the factor $(n+1)/n$ "as quick as lightning." Also, the distance already covered by the snail is being increased.

Let us look first at the individual 1 cm pieces the snail covers during the 1^{st}, 2^{nd}, . . . , n^{th} second, in particular at the stretches that have an effect on these pieces.

The 1 cm piece of the first second is at the transition $1^- \rightarrow 1^+$ being stretched to its double length, at the transition $2^- \rightarrow 2^+$ to the 3/2-fold length, . . ., in general at the transition $n^- \rightarrow n^+$ to the $(n+1)/n$-fold. So at the time n^+ it has the length (in cm):

$$1 \times \frac{2}{1} \times \frac{3}{2} \times \frac{4}{3} \times \ldots \times \frac{n+1}{n} = \frac{n+1}{1}.$$

The 1 cm piece along which the snail creeps during the second second is in the transition $1^- \rightarrow 1^+$ not yet stretched as a covered distance; it is being stretched only at the transition $2^- \rightarrow 2^+$ to the 3/2-fold length, at the transition $3^- \rightarrow 3^+$ to the 4/3-fold length, . . . , in general at the transition $n^- \rightarrow n^+$ to the $(n+1)/n$-fold length. It therefore has at the time n^+ the length (in cm):

$$1 \times \frac{3}{2} \times \frac{4}{3} \times \ldots \times \frac{n+1}{n} = \frac{n+1}{2}.$$

The 1 cm piece along which the snail creeps during the third second is being stretched only by the third stretching onward, for the length of this piece at the time n^+ we get analogously (in cm):

$$1\times\frac{4}{3}\times\frac{5}{4}\times\ldots\times\frac{n+1}{n}=\frac{n+1}{3}.$$

Finally we look at the 1 cm piece that is covered in the n^{th} second, which is stretched exclusively at the transition $n^- \to n^+$ with the stretching factor $(n + 1)/n$. So this piece has at the time n^+ a length of $(n + 1)/n$ cm.

The sum of all these stretched 1 cm pieces indicates the total distance of the snail from the starting point A. This distance is therefore at the time n^+:

$$\frac{n+1}{1}+\frac{n+1}{2}+\frac{n+1}{3}+\ldots+\frac{n+1}{n}=(n+1)\times\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}\right).$$

The band itself has, at the time n^+, a length of $n + 1$ km = $(n + 1) \times 10^5$ cm. The difference is the distance which the snail still has to cover:

$$(n+1)\times\left(10^5-1-\frac{1}{2}-\frac{1}{3}-\ldots-\frac{1}{n}\right)$$

The negative fractions in the second parentheses are exactly the terms of the harmonic series, from which one knows that the series is divergent—that is, it grows beyond all limits. The numerical value of the series

$$\sum_{i=1}^{n}\frac{1}{i}$$

is for a very big n therefore also bigger than 10^5, and the value of the expression in the second parentheses is therefore negative; but if the distance that remains to be covered is no longer positive, then this means that the end has been reached already. It takes, of course, a very long time for this to happen. To get at least a very rough approximation of the order of magnitude of the necessary period of time, a formula for the approximation for a very large n could help. The exact foundation of the following limit-statement (Euler-Mascheroni's constant):

$$\lim n\to\infty\left[\ln n-\sum_{i=1}^{n}\frac{1}{i}\right]=0.57721\ldots$$

will hardly be possible at school, but the teacher could just state it, without proof, to give an idea of the necessary period. For large n, we therefore have:

$$\ln n-\sum_{i=1}^{n}\frac{1}{i}=0.57721\ldots\Leftrightarrow\sum_{i=1}^{n}\frac{1}{i}=\ln n-0.57721$$

From the inequality $\ln n - 0.57721 > 10^5$, it follows immediately for n

$$n > e^{10^5 + 0.57721} \approx 1.78 \times 10^{43{,}429.45} \approx 5 \times 10^{43{,}429}.$$

This is the time in seconds the walking-tour takes the snail. But even expressed in years the number does not look much smaller—it would be approximately $1.6 \times 10^{43{,}422}$ years, an unimaginably huge number and an even more unimaginable period of time (according to today's knowledge, about 2×10^{10} years have passed since the Big Bang)!

References

Baron, G., & Windischbacher, E. (1990). *Österreichische Mathematikolympiaden 1970-1989: Aufgaben und Lösungen.* Innsbruck: Universitätsverlag Wagner.

Dürschlag, S. (1983). Problemlösen und Kreativität im Mathematikunterricht—Konzeption und Tendenzen in der polnischen Didaktikliteratur. *Der Mathematikunterricht, 29*(3), 46-70.

Engel, H. (Ed.). (1979). Internationale Mathematik-Olympiade. *Der Mathematikunterricht, 25*(1).

Haussman, K. (1986). Iteratives versus rekursives Denken beim Problemlösen im Mathematikunterricht. *Mathematica Didactica, 9,* 61-74.

Heinze, G. (1993). Förderung mathematisch-naturwissenschaftlich begabter Schüler in Sachsen am Beispiel der Mathematikausbildung am Gymnasium Dresden Blasewitz. *Didaktik der Mathematik, 21*(2), 140-153.

Kranzer, W. (1989). *So interessant ist Mathematik.* Köln: Aulis-Deubner.

Laub, J., Hruby, E., Reichel, H.-C., Litschauer, D., & Gross, H. (1985-1988). *Mathematik Arbeitsbuch* (4 vols.). Wien: Hölder-Pichler-Tempsky.

Polya, G. (1954). *Mathematics and plausible reasoning* (2 vols.). Princeton, NJ: Princeton University Press.

Polya, G. (1980). *Mathematical discovery* (2 vols.). New York: John Wiley.

Reichel, H.-C. (1991). Sprachschulung und Spracheinsatz im Mathematikunterricht. In H. Postel, A Kirsch, & W. Blum (Eds.), *Mathematik Lehren und Lernen* (pp. 156-170). Hannover: Schrödel.

Reichel, H.-C., Müller, R., Hanisch, G., & Laub, J. (1989-1992). *Lehrbuch der Mathematik* (Vols. 5-8). Wien: Hölder-Pichler-Tempsky.

Sell, R. (1988). *Angewandtes Problemlöseverhalten—Denken und Handeln in komplexen Zusammenhängen.* Berlin: Springer.

Stowasser, R., & Mohry, B. (1977). Rekursive Verfahren aus der Mathematikgeschichte für den Unterricht. *Der Mathematikunterricht, 23*(1), 5-41.

Zimmermann, B. (1983). Problemlösen als eine Leitidee für den Mathematikunterricht—Ein Bericht über neuere amerikanische Beiträge. *Der Mathematikunterricht, 29*(3), 5-45.

11

Another View of Combinatorics (or Counting Without Really Counting)

STEPHEN E. MORESH

Part One

The field of combinatorics has always been very interesting to both the mathematician and the lay person. Problems in this area are often related to the real world and are usually a great deal of fun. Unfortunately, students studying this discipline, whether at the secondary or university level, often complain that the problems are very difficult. When asked why they find it so hard, the usual response is that the type of reasoning involved in combinatorics is different from that in other mathematics courses they have taken. In fact, some students, trying to classify problems by "types," often give up in distress, after ending up with scores of different types of problems.

If you ended up with so many types of problems, you might want to give up, too; however, the analysis of most problems in combinatorics can be narrowed down to two basic categories. This chapter will focus on combinatoric problems a little differently, presenting two processes that can be used to solve most problems. The only assumption being made here is that the reader has had some experience (limited as it may be) with permutations, combinations, calculating ${}_nP_r$, ${}_nC_r$, and doing simple probability problems.

Combinatorics is usually introduced through a simple problems such as the following:

COMB In how many ways can a man choose a shirt and a tie if he has three shirts and four ties?

It is easy to show that there are 12 possibilities.

Traditionally, teachers demonstrate two approaches: The first is to draw a type of diagram called a "tree diagram":

⑥ ⑪

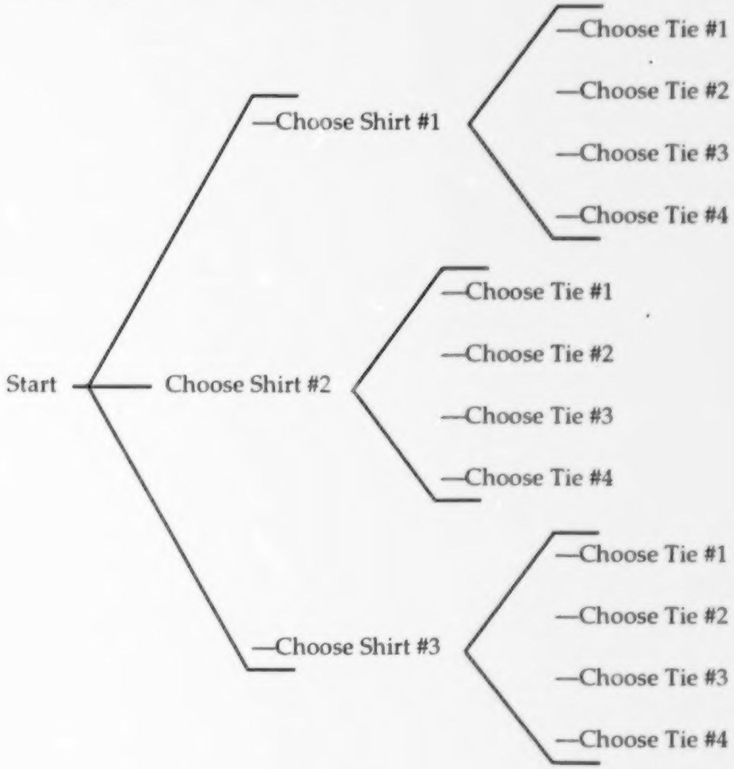

Start
—Choose Shirt #1
Choose Shirt #2
—Choose Shirt #3
—Choose Tie #1
—Choose Tie #2
—Choose Tie #3
—Choose Tie #4
—Choose Tie #1
—Choose Tie #2
—Choose Tie #3
—Choose Tie #4
—Choose Tie #1
—Choose Tie #2
—Choose Tie #3
—Choose Tie #4

Figure 11.1

⑥ It is clear that since this "tree" has 12 different paths, each one representing a different outfit, the answer to the problem is 12. And so we have *Combinatorics Problem-Solving Technique 1*: Draw a tree diagram.

The second approach teachers demonstrate is to use the fundamental counting principle. Although drawing a diagram is often an excellent problem-solving strategy, it is clear that with more shirts and ties, the diagram would become unmanageable.

At this point, the meaning of permutation and the fundamental counting principle are usually introduced. A *permutation* of a number of objects is any arrangement of these objects in a definite order. By using the shirt-and-tie and similar problems, we usually demonstrate the fundamental counting principle, also called the multiplication principle: If one thing can be done in n_1 ways, and after it is done another thing can be done in n_2 ways, and after it is done another thing can be done in n_3 ways, . . . , and after it is done another thing can be done in n_k ways, then the *k* operations can be performed in $n_1 \times n_2 \times n_3 \ldots \times n_k$ ways.

COMB

COMB

This gives the result 3 × 4 = 12 possible outfits in the original problem. We often teach this by having students fill in two blanks with the amount of ways each of the two stages can be performed, and then multiplying by the above principle.

Although this is quite elementary, the fundamental counting principle really summarizes a process that is vital to problem solving in combinatorics, although it is rarely spelled out that way. The key word in the fundamental counting principle is "and." The key to using this principle in a problem is recognizing an "AND process," a situation in which something is done AND then something else is done.

The shirt-and-tie problem can be interpreted as being an example of an AND process, since the man first chooses a shirt and then, afterwards, chooses a tie. This is why multiplying the number of ways a shirt can be chosen by the number of ways a tie can be chosen gives the correct result. An AND process is usually a hint that we must multiply to find the total number of ways things can be done. We state this as *Combinatorics Problem-Solving Technique 2*: For any situation that can be interpreted as an AND process, use the fundamental counting principle and multiply to find the total number of possibilities.

LOGIC

Although many problems obviously use an AND process, the use of such a process in other problems is not so obvious. In fact, as the following problem illustrates, the usual interpretation would not be to use an AND process. Viewing it with such a process will yield a different view of combinatoric problems and may make similar problems a great deal easier to solve.

> Mr. Appel likes to eat lunch in a particular Italian restaurant. He often has pizza, one of the restaurant's specialties. The menu indicates that he can have a plain slice of cheese pizza or any or all of the following toppings: pepper, onions, sausage, mushrooms, broccoli, or anchovies. For example, one day he has a slice of pizza with sausage and broccoli on top. If he decides to have a different type of slice each day, in how many days will he have sampled every possible type of pizza?

COMB

⑪ *Traditional Approach.* Since Mr. Appel is selecting toppings, order does not count. Therefore, we can consider this to be a problem involving a combination, a selection of objects without regard to order.

COMB Mr. Appel has 6 toppings from which he can choose 0, 1, 2, . . . , 6 toppings for his pizza. Thus, there is a total of $1 + 6 + 15 + 20 + 15 + 6 + 1 = 64$ different pizzas, so the answer to the original problem is 64 days. Most people are amazed at this, since they intuitively expect the answer to be at least several hundred. (Note that the symbol , in some texts written as $_nC_r$ or $C(n, r)$, or $\binom{n}{r}$ is used to mean the number of ways r items can be selected from n items.)

③ *AND process approach.* Although "and" does not even appear in the original problem, an AND process is implied. If we can find a way of rethinking the problem to find such an AND process, it will greatly simplify the solution.

Picture Mr. Appel sitting in the restaurant, deciding on the type of pizza he will have today. He starts by looking at the different toppings and asks himself, Should I have pepper? AND Should I have onions? AND Should I have sausages? AND Should I have mushrooms? AND Should I have broccoli? AND Should I have anchovies?

GEN For each question, he has to decide whether or not to include the specific topping. For example, for the first question (Should I have pepper?), the answer is Yes or No. This gives 2 possibilities for the first stage of the process. After deciding about pepper, he goes on to the next stage (Should I have onions?), and after deciding about onions, he decides about sausages, then mushrooms, and so forth. This can be summarized as a six-stage AND process. Since there are 2 choices at each stage, there are $2^6 = 64$ possible pizzas. Note that this includes the plain cheese pizza, in which the decision at each stage was not to take the given topping. A major advantage of this approach is that it can greatly cut down the amount of calculations. If the problem had originally included 15 different toppings for the pizza, the traditional approach would be quite messy indeed. This approach would yield $2^{15} = 32{,}768$ different toppings immediately. An AND process could be thought of as any given situation in which there is a sequence of stages in order to accomplish what is being described. This problem is equivalent to finding the number of subsets there are for a given set.

Other Problems Applying AND Processes

COMB

> In many places, the custom is to give a "tip" for good service, whether in a restaurant, where a waiter has taken your order, at the airport, where a porter has carried your suitcase, or in a building, where the doorman helped you carry some packages. In the United States and other countries, the tip is not an automatic part of a bill but given at the

> discretion of the person being served. Ms. Smith decided to give her doorman a little tip for helping her with a package. If her purse has only a quarter, a half dollar, a silver dollar, and a five-dollar bill, how many tips are possible?

Here the money denominations replace the pizza toppings. Using the approach of the pizza problem, Ms. Smith looks at her quarter and decides whether or not to give it to the doorman, giving two possibilities, AND then does the same, one by one, for each coin and bill. Thus, there would appear to be $2^4 = 16$ possible tips. This includes the case of her being very generous and giving *everything* to the doorman; however, it also includes the extreme opposite case of giving nothing at all. Since the problem states that she definitely gives a tip, we must exclude that case from the 16, giving an answer of 15 possible tips.

Note that we had to look back to see if any special or extreme cases had to be handled. This is very important, since it may be implied in a problem. Sometimes it is not clear at all, and two answers must therefore be given to take into account the possibility of special cases. We state this as *Combinatorics Problem-Solving Technique 3*: Look back to check for special cases. ⑤

> A porter helped someone retrieve luggage at the airport. How many possible tips can the porter receive, if the person has four pennies, a nickel, a dime, and six quarters?

Why is this more difficult than the previous problem? We have to be extra careful not to overcount the number of possibilities. For example, a 6-cent tip is just one possibility, but it can be achieved in four ways. Use the nickel and the first penny, the nickel and the second penny, and so on. Just shifting our thinking slightly will enable us to incorporate an AND process into the picture. Some types of coins have multiple possibilities! Rather than looking at each coin and determining whether or not to take it, we can envision the process as looking at the type of coin, or denomination, and seeing the number of possibilities.

First, decide about the pennies: Take 0, 1, 2, 3, or all 4 of them → 5 possibilities. ⑪ AND once you're finished, decide about the nickel: Take it or leave it → 2 possibilities. AND once you're finished, decide about the dime: Take it or leave it → 2 possibilities. AND once you're finished, decide about the quarters: Take anywhere from 0 to 6 quarters → 7 possibilities.

This would normally give $5 \times 2 \times 2 \times 7 = 140$ possible tips.

We must now go back to see if there are special cases. The extreme cases should always be checked first. The case in which the person gives the porter every ⑤ coin is certainly allowable. The other extreme, in which no pennies, no nickels, no dimes, and no quarters are taken is debatable! One may argue that the problem implies that the porter definitely gets a tip. With this interpretation, we must eliminate the case of no coins. Here the answer would be 140 – 1 = 139 possible tips.

Someone else may argue that giving no coins is a tip as well—a tip of $0.00! Including this would leave the answer at 140 tips. Although most people would not accept this argument, since the problem does not say that the porter definitely received a tip, we can accept this answer as well!

This last example certainly demonstrates that when we test this concept, we should make sure that each question is crystal clear. Rewording the problem slightly can clarify whether the answer should be 139 or 140.

> An assistant helped Ms. Streett retrieve her coat after she ate in a fancy restaurant. When she looked in her purse, all she had was a twenty-dollar bill, two dimes, two quarters, and a nickel. She decided definitely to give a tip, but not the twenty-dollar bill. In how many ways may she give the tip?

Using the procedures of the previous problems, we would multiply 3 × 3 × 2 = 18. Remembering that she will definitely give a tip, we deduct the case of no tip, giving an answer of 17.

⑪ Listing the 17 outcomes may shed further light on this problem (we use *N* to stand for nickel, *D* for dime, *Q* for quarter). The possible tips, then, are:

⑯

One-coin tips:	*N, D, Q*
Two-coin tips:	*ND, NQ, DD, QQ, DQ*
Three-coin tips:	*DDN, DDQ, DQN, QQN, QQD*
Four-coin tips:	*DDQQ, DDQN, QQDN*
Five-coin tips:	*QQDDN*

There are clearly 17 different *ways* of tipping, which the problem asked for. Had the question asked *how many different tips* are possible, the answer would be different. Listing the values of the tips (in cents), we get:

One-coin tips:	5, 10, 25
Two-coin tips:	15, 30, 20, 50, 35
Three-coin tips:	25, 45, 40, 55, 60
Four-coin tips:	70, 50, 65
Five-coin tips:	75

Note that 25- and 50-cent tips appear twice. Why does this happen? In the original problem, combinations of some of the coins add up to other coins. Specifically, the two dimes and a nickel add up to a quarter. Two dimes and a nickel are a different way of tipping from leaving one quarter; however, they have the same total value. In the previous tipping problems, this difficulty did not occur. In this problem, because of the double occurrence of 25- and 50-cent tips, we must subtract 2 from the 17 tips listed above, thus giving us 15.

This last example illustrates that we must read all problems very carefully. Counting problems are infamous for the number of nuances that can occur. And so we come to *Combinatorics Problem-Solving Technique 4*: Read the problem very carefully until you understand exactly what is being asked, and be sure to clear up any ambiguity before attempting to solve the problem.

COMB

If we were going to be a little more realistic, we would add a condition that a certain minimum tip was being considered. Other variations and approaches to the last tipping problem can be considered.

Finding the number of committees that can be formed is a popular type of question for students studying combinations. The following problem illustrates how AND processes can make a difficult committee problem much easier.

> A club consists of 10 married couples. In how many ways can a committee be formed consisting of 3 people if a husband and wife cannot serve together on the committee?

COMB

Method 1. Picture the committee-forming process: This is basically an AND process. First, choose the 3 couples to be represented on this committee, AND then pick 1 representative from each couple. This gives $\binom{10}{3} \cdot 2 \cdot 2 \cdot 2 = 120 \cdot 8 = 960$ committees.

Method 2. This method also uses an AND process. Here, however, we do the complementary problem; that is, we will make sure that at least 1 husband-and-wife team are represented. Then subtract this number from the number of committees possible with no restriction.

③

With no restriction, there are $\binom{20}{3} = 1{,}140$ committees.

For the complementary problem, we must ensure that at least 1 husband serves together with his wife. Picture choosing a committee by handling this part first and then choosing the remaining committee members. There are 10 ways of choosing a husband together with his wife, since there are 10 couples. After choosing the couple, we need one more person to complete the committee. There are 18 ways of doing this. Thus, the complementary problem has 10 × 18 = 180 committees, giving us 1,140 – 180 = 960 committees, where husband and wife do not sit together.

③ Although the first method was easier here, the second method was presented since it will be the easier approach in other problems. In addition, it gives us *Combinatorics Problem-Solving Technique 5*: Do the complementary problem and subtract this answer from the total number of outcomes in general.

Exercises

GEN For each problem, *(a)* interpret the problem using an AND process; *(b)* solve the problem; and *(c)* if possible, find another (more traditional) solution.

COMB 1. A certain locksmith makes keys using any one of 15 brands of keys. Each brand uses a different length and thickness. No matter what brand is chosen, there are 8 different positions from which metal can be removed or left alone. For the first 5 positions, there are 3 different depths at which the metal can be removed. For the remaining positions, there are 4 different depths. How many different keys are possible if metal must be removed in at least one position?

COMB 2. How many subsets does a set of 15 elements have?

COMB 3. Many different kinds of desserts have been put out for people to choose from at a party. Judy is watching her weight and decides to have at least one piece of fruit and no other type of dessert. If the fruit bowl contains exactly 6 apples, 5 plums, 3 oranges, 4 bananas, and 5 pears, how many different combinations of desserts can Judy have?

COMB 4. In how many ways can Peggy make a selection of one or more new fish for her tank from 6 identical coral fish, 7 identical angel fish, and 3 identical blue pullers?

COMB 5. A postal carrier is late and wants to hurry up to go home. He decides to just put any letter into any mailbox. If there are 12 letters and 20 mailboxes, in how many ways can the letters be placed if *(a)* each letter is to go in a different mailbox; *(b)* any number of letters may be placed in a mailbox?

COMB 6. In how many ways can the letters in exercise 5 be placed if there are 20 letters and 12 mailboxes?

COMB 7. Problems 5 and 6 above can be generalized. In how many ways can a set of m items be matched with a set of n items if $m < n$ and *(a)* each of the m items can be matched with at most one of the n items; *(b)* there is no restriction on the number of items with which any items can be matched?

COMB 8. There are 5 different algebra books, 4 different biology books, and 6 different calculus books. How many selections of books can Roslyn make if *(a)* she decides to take at least one book; *(b)* she decides she must include a specific algebra book; *(c)* she decides she must include an algebra book but it could be any one of them; *(d)* she decides she must take at least one algebra and at least one biology book; *(e)* she must have at least one of each type of book?

9. *(a)* In how many ways can a triangle be named using letters of the alphabet? *(b)* In how many ways can a triangle be named using letters of the 24-letter Greek alphabet? *(c)* In how many ways can a quadrilateral be named using letters of the alphabet? *(d)* In how many ways can an *n*-gon be named using letters of the alphabet? What restriction is there on the value of *n?* GEOM
10. After breakfast in a diner, Mrs. Stux looked for change in her purse to give the waiter a tip. She had 6 dimes, 2 quarters, 4 dimes, and 3 pennies. How many tips can she give if *(a)* she decided to definitely give a tip; *(b)* she did not like the service so she will give at most 5 coins; *(c)* she did not like the service so she will give at most 5 coins but will definitely give a tip; *(d)* she did like the service so she will give at least five coins; *(e)* she liked the service and intends to give at least 75 cents. COMB

Part Two

How is the following problem different from previous ones that we have considered?

Brenda is going on vacation. She likes to read and decides to borrow 2 or 3 books from the library. Suppose also that she has narrowed her selection down to 6 books. In how many ways can she make her selection? COMB

The major difference here is that we have what we will refer to as an "OR process." She can take 2 OR 3 books from the library. She cannot take exactly 2 AND exactly 3 books at the same time. Thus, there are different cases to consider.

In this problem, the most important word is "or." If you can recognize that a problem involves an OR process, it may greatly simplify its solution. An OR process is any situation in which either one thing will be done or another thing will be done, but not both of them. LOGIC

In order to solve the vacation problem, analyze each case as if that case were the entire problem:

Case 1: Brenda chooses exactly 2 books. She can do this in $\binom{6}{2} = 15$ ways. ⑪

Case 2: Brenda chooses exactly 3 books. She can do this in $\binom{6}{3} = 20$ ways.

These two cases cannot occur simultaneously (she will choose either exactly 2 books OR exactly 3 books but not both, since both events cannot happen at the same time). This leads to the following definition: Two events are said to be "mutually exclusive" or "disjoint" events if they cannot possibly happen at the same time. COMB

The events described in cases 1 and 2 are obviously mutually exclusive events. Since each of the 15 outcomes in case 1 contains 2 books and each of the 20 outcomes in case 2 contains 3 books, there is no overlap. In other words, the 15 outcomes in case 1 are all different from the 20 outcomes in case 2. Thus, the total number of selections that Brenda can make is 15 + 20 = 35.

We now have *Combinatorics Problem-Solving Technique 6*: For any situation that can be interpreted as an OR process, handle each possible case separately and then add to find the total number of possibilities.

A hint that an OR process is involved is the presence of mutually exclusive events. Thus, there is another way to state *Combinatorics Problem-Solving Technique 6:* If two events are mutually exclusive, and the first event can occur in x ways and the second can occur in y ways, then the total number of ways that the first event or the second event can occur is the sum, $x + y$.

Very often, problems involve an OR process in a way that is not obvious. Let us focus on some of these.

ARITH How many whole numbers less than 1,000 can be created if the digits must come from the set {3, 5, 6, 7, 9}?

If the number is to be less than 1,000, there are three mutually exclusive cases to handle:

(11)

Case 1: The number created has one digit. There are 5 possibilities here.

Case 2: The number created has two digits. There are $5 \times 5 = 25$ possibilities here. (Note that repetition of digits is permitted, since the original problem had no prohibition against this.)

Case 3: The number created has three digits. There are $5 \times 5 \times 5 = 125$ possibilities here.

Since these three cases are mutually exclusive, we add the number of ways in which each case can be accomplished, giving 5 + 25 + 125 = 155 numbers.

A problem that has interested state and federal officials has been that of creating enough license plates. When a region runs out of license plate numbers, new rules are often formulated in order to create enough new numbers to allow for the growing number of vehicles.

> A certain state has license plates which consist of three letters of the alphabet followed by a three-digit number greater than or equal to 100. It also produces, for a fee, "vanity" plates, such as "DAVID," "MARY 1," or "NURSES." The Bureau of Motor Vehicles, in order to provide for more cars, trucks, and buses, produced new license plates consisting of three letters followed by a four-digit number greater than or equal to 1,000. How many nonspecial licenses can this state produce?

COMB

In the above problem, there is an OR process implied, since there are three cases—special licenses, those that consist of six characters using the old rules, and those with seven characters using the new rules. Since we are not interested in the special licenses, we need only focus on the last two cases.

⑯

Case 1: Six-character licenses are formed using an AND process, since a character is chosen first, and then another, and another, and then three digits. There are 26 possibilities for the first letter-character, 26 for the second, and 26 for the third; for the fourth character (a non-zero digit) there are 9 possibilities, 10 possibilities for the fifth character (a digit), and 10 for the sixth (also a digit). This gives $26^3 \times 9 \times 10 \times 10 = 15{,}818{,}400$ possible license plates.

Case 2: Seven-character licenses are also formed using an AND process. The possibilities are the same as those in the six-character place, with the addition of 10 possibilities for the seventh-character (a digit). This gives $26^3 \times 9 \times 10 \times 10 \times 10 \times 10 = 158{,}184{,}000$ possible license plates.

Thus, the total number of licenses possible is 174,002,400. Had the state used the last rule in the first place, it probably would not have needed a change, since this rule alone provides for many more vehicles than it would ever have. In actuality, when states figure out how many licenses are possible, there are many special cases that must be excluded from consideration. For example, in some states the following three-letter combinations are not allowed for obvious reasons: KKK, LSD, SEX, JEW, and so forth.

License plate problems are very typical of many OR process problems in that each case involves an AND process. Each case is therefore handled using multiplication, whereas the final result is handled by adding the amounts given by the various cases. We state this as *Combinatorics Problem-Solving Technique 7*: Each case of an OR process can often be interpreted as an AND process.

COMB

LOGIC

Exercises

ARITH
NUMB

1. How many whole numbers less than 1,000 can be created if the digits must come from a set of 8 different non-zero one-digit numbers?

ARITH

2. How many whole numbers less than 10,000 can be created if the digits must come from a set of *n* different non-zero one-digit numbers?

COMB

3. How many license plates would a state have using the rule that the license have *(a)* four letters followed by a two-digit number; *(b)* four letters followed by a two-digit number or a two-digit number followed by four letters; *(c)* six characters which can be digits or letters?

COMB

4. How many licenses are possible if a state allows licenses to contain three letters followed by three digits or three digits followed by three letters?

COMB

5. A state creates license plate numbers by using as a general rule three letters followed by three digits. Every possible license number has been used! Since it is expensive to create new licenses for everybody, it is probably cheaper to add a new character for new licenses. Explain how to generate new licenses so that the maximum number of plates is available.

COMB

6. How many different selections of at least one book can Erica make from 7 different books? *(a)* Do this problem using an AND process. *(b)* Do this problem using an OR process. For this problem, which process produces an easier solution?

COMB

7. A committee of at least 2 people is to be made from 8 people. How many different committees are possible?

COMB

8. Radio stations in the United States can have three or four call letters. East of the Mississippi River, they start with the letter *W*, while west of it they start with the letter *K*. How many different radio stations are possible using these rules?

COMB

9. Tony's Tattoo Parlor offers the following designs: a heart with a dagger through it, an eagle, a rose, a panther, a skull, or a design with MOM, DAD, your name, or your girlfriend's name on it. No one gets the same tattoo twice. *(a)* Yesterday, Bob was seen going into Tony's. No one knows how many tattoos he got. He may have even lost his nerve and gotten none! How many different selections of tattoos could Bob now have? *(b)* Suppose you heard that Bob definitely got a design with "MOM" put on his right arm. How many different selections of tattoos could Bob now have? *(c)* Suppose you heard that Bob definitely got five tattoos including one with his girlfriend's name on it. How many selections of tattoos can Bob now have?

COMB

10. A certain diagram has 20 points in a plane, no 3 of which are collinear. How many different *(a)* triangles can be connected? *(b)* quadrilaterals can be connected? *(c)* hexagons be connected? *(d)* polygons can be connected? *(e)* Redo part *d* another way.

Part Three

In this section we consider some applications of the previous two sections to solving problems in probability.

When we dealt with AND processes, we saw that we have to multiply to find the total number of ways two operations could be done. The most natural question to ask is whether or not we can multiply probabilities when an AND process is involved. To investigate this, we consider the following problem:

LOGIC

> An urn contains three red marbles and one blue marble. If two marbles are chosen randomly, find the probability they are both red if *(a)* the first marble is returned to the urn before the second one is chosen; *(b)* the first marble is not returned to the urn before the second one is chosen.

PROB

Both parts of this problem involve AND processes. The problem turns out to be more difficult than it appears! Two reds can occur in many ways. Both parts of the problem are very important, since replacement problems illustrate an important problem-solving technique related to AND processes.

What are the different problem-solving strategies used to solve the two parts of this problem?

First, to simplify the problem, we envision each of the red marbles having the number 1, 2, or 3 on it. We will let R_1 stand for the first red marble, R_2 the second, R_3 the third, and B the blue marble. Thus, the sample space is $\{R_1R_1, R_1R_2, R_1R_3, R_1B, R_2R_1, R_2R_2, R_2R_3, R_2B, R_3R_1, R_3R_2, R_3R_3, R_3B, BR_1, BR_2, BR_3, BB\}$. Of the 16 possibilities, 9 have red in both cases (R_1R_2, for example). Thus, the desired probability is 9/16. Note the following probabilities associated with this problem: P(red the first time) = 3/4, P(red the second time) = 3/4. P(red both times) turns out to be the product of the two probabilities. ⑪

In the second strategy, the sample space is different, since certain cases such as R_1R_1, for example, cannot be included. What happened in the first event affects what happens the second time. The sample space is therefore $\{R_1R_2, R_1R_3, R_1B, R_2R_1, R_2R_3, R_2B, R_3R_1, R_3R_2, R_3B, BR_1, BR_2, BR_3\}$. Of the 12 possibilities, 6 have red in both cases. Thus, the desired probability is 6/12. Here the associated probabilities are: P(red the first time) = 3/4. P(red the second time) = 9/12 = 3/4. P(red both times) does not turn out to be the product of the 2 probabilities. ⑪

What is different about these cases that allows us to multiply probabilities the first time but not the second time? The difference is that in the first case, the first outcome has no effect on the second outcome. In the second case, what happens the first time *does* affect the second. Here, for example, blue for the first selection disallows blue in the second selection.

PROB

If an AND process is involved in a probability problem, you must ask yourself, Does the first event affect the second event?

When the result of one event has *no* effect on the result of a second event, the two events are said to be independent. As you saw in the previous problem, when an AND process problem involves two independent events, we can multiply probabilities. To make this even simpler, so you would not have to make subjective decisions about how events are related, mathematicians define independent events more mechanically: Two events A and B are said to be independent events if $P(A \text{ AND } B) = P(A \wedge B) = P(A) \times P(B)$. We have, then, *Combinatorics Problem-Solving Technique 8*: For a problem involving an AND process and two independent events A and B, $P(A \wedge B) = P(A) \times P(B)$.

LOGIC

Beware! When asked for $P(A \wedge B)$, do not automatically multiply two probabilities. If A and B are *definitely* independent events, go ahead. If they are not, you may have to resort to other techniques, such as drawing a tree diagram and so on. Often, the topic of conditional probability (not considered here) will apply in such examples.

What is the most natural question to ask at this point? When we dealt with OR processes, we *added* to find the total number of ways the first or the second could be done. Can we add probabilities when an OR process is involved? Consider the following problem:

PROB

Maile randomly pulls a card from a standard deck of 52. *(a)* What is the probability that she pulls a seven or an ace? *(b)* What is the probability that she pulls a seven or a red card?

Both parts of this problem involve OR processes. They will illustrate problem-solving strategies needed to deal with probability problems involving OR processes.

In part *a*, $P(\text{Maile pulls a seven}) = 4/52$; $P(\text{she pulls an ace}) = 4/52$. There are 8 cards that are sevens or aces, so that $P(\text{she pulls a seven OR an ace}) = 8/52$. Note the additive nature of the probabilities involved. In other words, $P(A \text{ OR } B) = P(A \vee B) = P(A) + P(B)$.

In part *b*, $P(\text{she pulls a seven}) = 4/52$; $P(\text{she pulls a red card}) = 26/52$. There are 28 cards that are sevens or red, so that $P(\text{she pulls a seven OR a red card}) = 28/52$. Note that, in this case, the probabilities are not additive. In other words, $P(A \vee B) \neq P(A) + P(B)$. (Two cards, the seven of hearts and the seven of diamonds, would be counted twice if we just added the two probabilities.)

PROB

What made the probabilities in part *a* additive and the probabilities in part *b* non-additive? The key is mutual exclusivity, which we defined in part 2 of this chapter. In other words, $P(A \wedge B) = 0$.

How are mutually exclusive events different from independent events? Mutually exclusive events are the most highly dependent events, since $P(A \wedge B) = 0 \neq P(A) \times P(B)$.

If an OR process is involved in a probability problem, you must ask yourself, Can the two events possibly occur at the same time? This leads us to *Combinatorics Problem-Solving Technique 9*: For a problem involving an OR process and two mutually exclusive events A and B, $P(A \vee B) = P(A) + P(B)$. LOGIC

Case B, in which the two events were not mutually exclusive, comes under another category for which the formula $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$ can be used. This formula considers the fact that outcomes in both events are included twice—once in $P(A)$ and once in $P(B)$—so that the extra $P(A \wedge B)$ had to be subtracted. This more general formula can even be used for mutually exclusive events, since $P(A \wedge B) = 0$.

Exercises

For each problem, see if AND or OR processes can be used to find the given probability. Then find the desired probability.

1. A die is rolled 5 times. Find the probability that the number 2 shows up all 5 times. PROB
2. A nickel and a dime are tossed simultaneously. Find the probability that *(a)* both show heads; *(b)* both show tails; *(c)* the nickel lands tails and the dime lands tails; *(d)* one of them lands heads and the other lands tails. PROB
3. Helen tosses a coin 6 times. Find the probability that she gets heads the first 3 times and tails the last 3 times. PROB
4. The numbers 7, 8, 11, 12, and 15 are written on 5 slips of paper and put into a hat and stirred. Two numbers are picked, one after the other, without replacement. The two numbers are then noted. Find the probability that the sum of the numbers is odd. PROB
5. A box contains 7 marbles: 3 red and 4 blue. Two marbles are drawn, one after another. Find the probability that both are red if *(a)* the first marble is replaced before choosing the second one; *(b)* the first marble is not replaced. PROB
6. Roz bought 12 cups of yogurt. Unfortunately, 4 of them are spoiled. Find the probability that the first 3 cups she chooses turn out to be spoiled. PROB
7. Blake is a senior in high school. He estimates the probability of being accepted at college *A* to be 0.75 and the probability of being accepted at college *B* to be 0.5. He then multiplies, and claims that the probability of being accepted at both colleges is 0.375. Explain why Blake's argument is or is not correct. PROB

Part Four

In many combinatorics and probability problems, certain events are described as "success" or "failure." For example, in finding the probability that a card chosen randomly from a standard deck is a picture card, we often give the answer 12/52, because there are 12 possible "successes," meaning that there are 12 picture cards out of the 52 cards in the deck. We often give our beginning students the following definition: The probability of an event is equal to the number of ways a success can occur divided by the total number of outcomes possible.

We really must be careful with the use of the terms "success" and "failure." For example, a research scientist doing an experiment on rats may consider "getting cancer" a success. In real life, no one in his right mind would consider getting cancer a success. Therefore, our interpretation of success and failure in probability is quite different from the interpretation in real life.

What do we mean when we use the term "success" in combinatorics and probability problems? "Success" merely means meeting a certain requirement (whether desirable or not).

In most problems, we use "success" or "failure" to mean meeting or not meeting the conditions of the problem. This means that, in many trials, we can think of the problem as having two possible outcomes: success and failure. This is used in many types of problems, especially those involving the repeating of trials that are independent of one another. About three hundred years ago, the first person to study this was the famous mathematician James Bernoulli (1654-1705). In his honor, we make the following definition: An experiment is called a "Bernoulli trial" if the following three conditions are met:

1. We must be able to interpret the results of each trial as one of two outcomes, one called "success" (S), the other called "failure" (F).
2. The result of each trial must be independent of previous (and future) trials.
3. The probability of success, denoted by p, and the probability of failure, denoted by $q = 1 - p$, cannot change from trial to trial. In other words, these two probabilities must remain constant.

Examples of Bernoulli Trials

PROB

1. If a coin is fair, and "landing heads" is considered to be a success, then $p = P(S) = 1/2$ and $q = P(F) = 1 - p$ is also 1/2. Each flip of the coin is independent of previous and future flips.
2. Suppose a problem involves the rolling of a fair die. Suppose also that in the problem, getting a four or a five will result in some extra monetary reward. Thus, we can consider getting a four or a five as a success. Thus $p = 2/6 = 1/3$ and $q = 4/6 = 2/3$.

3. If a problem involves picking cards randomly from a standard deck of cards and getting a seven is considered a success, this is another example of a Bernoulli trial, as long as the card you choose is returned to the deck each time. (Why?) Thus, $p = P(\text{picking a seven}) = 4/52 = 1/13$, and $q = P(\text{failure}) = P(\text{not picking a seven}) = 12/13$.
4. Someone has made a dartboard of his favorite politician. Suppose the politician's face composes 30% of the picture, and suppose also that the dart is just thrown randomly at the board. The probability of a hit is then 0.3 and a miss is 0.7. If the person does not throw randomly, these probabilities will be different. In fact, they may change so that we would not be dealing with Bernoulli trials; however, if each throw is done randomly, we can use the work of this section to find needed probabilities.

Bernoulli trials are independent of one another. As a result, we can use multiplication to find associated probabilities. Study the following examples and you will see some very interesting results.

More Examples

1. Glenda rolls a die 5 times, with "getting a number less than three" being considered a success. Find the probability that Glenda has 2 successes followed by a failure, a success, and a failure. What are the probabilities of success and failure on one throw of the die? Note that $p = 2/6 = 1/3$ and $q = 2/3$. Since the results of the five rolls of the die are independent of one another, we can treat this problem using an AND process. Using combinatorics problem-solving technique 8, we can multiply to get the desired probability. This gives $P(SSFSF) = p \times p \times q \times p \times q = p^3q^2 = (1/3)^3 \times (2/3)^2 = 4/243$.
2. Josh, like Glenda, rolls a die 5 times, with "getting a number less than three" being considered a success. What is the probability that Josh has 2 failures followed by 3 successes? $P(FFSSS) = q^2p^3 = p^3q^2 = (1/3)^3 \times (2/3)^2 = 4/243$, the same as in the example above.
3. Adam, like Glenda and Josh before him, rolls a die 5 times, with "getting a number less than three" being considered a success. What is the probability that Adam has alternating successes and failures, starting with a success? $P(SFSFS) = p \times q \times p \times q \times p = p^3q^2 = (1/3)^3 \times (2/3)^2 = 4/243$, the same as before!

What do these three examples have in common? Of course, they are examples of Bernoulli trials with the same *p*'s and *q*'s. An AND process was implied in each example. In addition, each person had 5 trials, with 3 successes and 2 failures. To generalize these results, we ask the following: ②

PROB

In rolling a die 5 times with "getting a number less than three" being considered a success, what is the probability of getting exactly 3 successes?

Getting exactly 3 successes in 5 trials means getting exactly 2 failures as well. The three previous examples illustrate some specific ways of accomplishing this.

COMB

The current problem generalizes the procedure by combining Glenda's, Josh's, and Adam's results by allowing any three successes. Listing all the ways of getting exactly 2 failures and 3 successes in 5 trials yields the following 10 outcomes: *FFSSS, FSFSS, FSSFS, FSSSF, SFFSS, SFSFS, SFSSF, SSFFS, SSFSF, SSSFF.* Since all of these outcomes mutually exclude one another, and since each has the same probability of occurring, p^3q^2, we add these to get P(exactly 3 successes in 5 trials) = $10p^3q^2 = 40/243$.

Note that the key in doing this was to be able to count all the ways you can have 3 successes in 5 trials. How could we have found that there are 10 such possibilities? This is merely a straightforward combinations problem to compute $\binom{5}{3} = 10$.

Let us generalize these results even further and note any interesting patterns.

Table 11.1

PROB

Number of Successes	*Ways of Getting These Successes*	*Probability of This Number of Successes*
0	FFFFF (1 way)	$1q^5 = 1p^0q^5 = \binom{5}{0}p^0q^5$
1	SFFFF, FSFFF, FFSFF, FFFSF, FFFFS (5 ways)	$5p^1q^4 = \binom{5}{1}p^1q^4$
2	SSFFF, SFSFF, SFFSF, SFFFS, FSSFF, FSFSF, FSFFS, FFSSF, FFSFS, FFFSS (10 ways)	$10p^2q^3 = \binom{5}{2}p^2q^3$
3	SSSFF, SSFSF, SSFFS, SFSSF, SFSFS, SFFSS, FSSSF, FSSFS, FSFSS, FFSSS (10 ways)	$10p^3q^2 = \binom{5}{3}p^3q^2$
4	SSSSF, SSFS, SSFSS, SFSSS, FSSSS (5 ways)	$5p^4q^1 = \binom{5}{4}p^4q^1$
5	SSSSS (1 way)	$1p^5q^0 = \binom{5}{5}p^5q^0$

The above table will enable us to answer slightly more complicated problems. Try the next example before reading the solution.

> In rolling a die 5 times with getting a number less than three being considered a success, what is the probability of getting exactly 3 OR 4 successes?

The events getting exactly 3 successes and getting exactly 4 successes are mutually exclusive. As a result, we can use combinatorics problem-solving technique 9 and add to get the desired probability, giving $10p^3q^2 + 5p^4q^1 = 40/243 + 20/243 = 60/243 = 20/81$.

Study the table above. Where have you seen the entries in the last column before? Recall:

$$(p+q)^5 = \binom{5}{5}p^5q^0 + \binom{5}{4}p^4q^1 + \binom{5}{3}p^3q^2 + \binom{5}{2}p^2q^3 + \binom{5}{1}p^1q^4 + \binom{5}{0}p^0q^5.$$

Since the probabilities involving Bernoulli trials result in expressions that come from a binomial expansion, as above, we call such probabilities "binomial probabilities": In n Bernoulli trials, the probability of exactly k successes is $\binom{n}{k} \times p^kq^{n-k}$. ⑱

Now we come to *Combinatorics Problem-Solving Technique 10*: If a problem involves n repeated Bernoulli trials, use binomial probabilities.

Let us look at a few more problems, each one relying on binomial probabilities.

> At a certain convention with a large attendance, 40% of the attendees are from the United States. If 12 people are randomly chosen, find the probability that 6 of them are from the United States. PROB

What thought processes should be in our minds as we do this problem? The key hint is contained in the question, implying that choosing a person from the United States is a success and anyone else a failure. Since there are so many people in attendance, for all purposes, the probabilities of success and failure are not changing. In addition, the nationality of the person you choose is independent of previous choices. Thus, P(exactly 6 successes out of 12) = $\binom{12}{6}p^6q^6$. This gives 924(0.4)6(0.6)6 = 924(191,102,976/1,000,000,000,000) = 0.176570149 ≈ 0.18.

In problems with messy calculations, as often happens with binomial probabilities, it is a good idea to use a calculator to help simplify the problem. In the above, it is quite easy to simplify using a y^x button and having the calculator find 0.24^6 (or separately find 0.4^6 and 0.6^6 and then multiply). In the good old days, before calculators were widely available, it was not uncommon to find tables of ⑫ binomial probabilities in the appendices of probability and statistics books. This brings us to *Combinatorics Problem-Solving Technique 11*: Use a calculator to find most binomial probabilities.

> Deanna did not study for her test because she thought she could pass it without studying; to use the vernacular, she thought she could "wing it." If the test consists of 10 multiple-choice questions, each with 4 choices, and 70% is passing, what is the probability that she passes the test by just guessing?

Here, careful thought must take place before realizing that, although there are 4 choices each time, each outcome is either a success (getting the right answer) or a failure (getting the wrong answer). Her getting a question right or wrong is independent of what she did on the previous questions. Since she is just guessing, the probability of getting a correct answer to any question is $p = 1/4$ each time, and the probability of getting the wrong answer is $q = 3/4$. The only difficulty here is that we have to consider 4 mutually exclusive cases. After dealing with these cases, we then use combinatorics problem-solving technique 9 and add the 4 results.

⑪ Case 1: $\text{P(exactly 7 successes)} = \binom{10}{7} p^7 q^3 = 120(1/4)^7\,(3/4)^3$

$= 120(1/16{,}384)\,(27/64) = 3{,}240/1{,}048{,}576$

Case 2: $P(\text{exactly 8 successes}) = \binom{10}{8} p^8 q^2 = 45(1/4)^8\,(3/4)^2$

$= 45(1/65{,}536)\,(9/16) = 405/1{,}048{,}576.$

Case 3: $P(\text{exactly 9 successes}) = \binom{10}{9} p^9 q^1 = 10(1/4)^9\,(3/4)^1$

$= 10(1/262{,}144)\,(3/4) = 30/1{,}048{,}576.$

Case 4: $P(\text{exactly 10 successes}) = \binom{10}{10} p^{10} q^0 = 1(1/4)^{10}\,(3/4)^0$

$= 1(1{,}048{,}576)\,(1) = 1/1{,}048{,}576.$

Thus, the probability that Deanna passes her test is the sum, or 3,676/1,048,576, which is approximately 0.0034. Our only advice to poor Deanna is to study, study, study!

The beauty of this last problem is that it illustrates the importance of what we did in parts 1 and 2 of this chapter. Each case itself involves an AND process, while all the cases considered together give us an OR process. By viewing the problem this way, it really is not too difficult, except for doing some messy calculations.

We end this chapter with a problem that comes up every October in the United States and Canada.

Baseball is a favorite sport in several countries, especially the United States and Canada. At the end of the season, in October, the two best teams play each other to see which team can win 4 out of 7 games. Suppose that the New York Mets are to play the Toronto Blue Jays in a truly international World Series. Find the probability that the New York Mets win the entire series, assuming the two teams are equally good.

Before solving this problem, there are several considerations. Even though the World Series is supposedly a 7- game series, it may end in fewer games. If a team wins the first 4 games, or 4 out of the first 5 or 6 games, there is no need to continue on to a 7-game series. As a result, we have to be very careful. First, let us focus on a few related questions.

First, is the World Series an example of a Bernoulli trial?

PROB

1. There are two outcomes for each "trial": success if your favorite team wins, failure if it loses.
2. Are the results of one game independent of the results of previous or later games? This is very hard to establish. In actuality, the answer may be No! If a team wins, psychologically, the members may be so thrilled that they perform better in the next game. If a team loses, the members may be so depressed that they do worse in the next game. The reverse may even be true. The losing team in one game may become more aggressive and determined to win the next game, while the winning team may become more lax and lose the next game. Since there are so many of these factors that may or may not play a role, we will take the words of the problem literally and accept that in any game the two teams are equally matched. We therefore may be oversimplifying the problem by accepting independence.

3. We will accept the fact that the probabilities do not change, even though this may not be true as explained in (2), above. In fact, with slightly different lineups of players, the probabilities likely do change, if only slightly. Again, we may very well oversimplify matters by assuming that we have Bernoulli trials.

PROB

Second, what are the binomial probability parameters in this problem? Since the problem says the teams are equally matched, the probability of each team winning (and losing) is 1/2. Since the problem asks us to find the probability that the New York Mets win, we accept the Mets winning as success and the Blue Jays winning (the Mets losing) as failure. Thus $p = q = 1/2$.

The number of successes is obviously 4, since it takes that number of games to win. The biggest difficulty is the number of games in all. A World Series can be played in 4, 5, 6, or 7 games. If we were to make these separate cases, we run into the danger of counting several combinations more than once. For example, the winning combination *SSSS* is included in each of the other cases! *FFFF* is embedded in *SSSSF*, *SSSSFF*, *SSSSFFF*. We can simplify the work by realizing that the winner must win the *very last* game, and the only flexibility is in choosing the successes and failures among the games *before* the last one.

Method 1

⑪ Case 1. 4 games: $P(\text{winning all 4 games}) = p^4 = (1/2)^4 = 1/16$.

Case 2. 5 games: Since the Mets will definitely win the last game, we must find $P(\text{winning 3 of the first 4 games and winning the last game}) = \binom{4}{3}p^3qp = 4(1/32) = 1/8$.

Case 3. 6 games: We must find $P(\text{winning 3 of the first 5 games and winning the last game}) = \binom{5}{3}p^3q^3p = 10(1/64) = 5/32$.

Case 4. 7 games: We must find $P(\text{winning 3 of the first 6 games and winning the last game}) = \binom{6}{3}p^3q^3p = 20(1/128) = 5/32$.

Thus, the probability that the Mets win is the sum $1/16 + 1/8 + 5/32 + 5/32 = 1/2$. (Is this the answer you would expect?)

Method 2

It would have been easier had there always been 7 games. We can envision, for the sake of this problem, that the teams do play 7 games. Even if the Mets win the first 4 games, they could have continued giving such combinations as *SSSSFFF*, *SSSSSFS*, *SSSSFSS*, and so on. Thus, we can do the problem by finding the probability that in 7 games the Mets win at least 4 games: This gives *P*(Mets win 4 of 7 games) + . . . + P(Mets win all 7 games) = $\binom{7}{4}p^4q^3 + \binom{7}{5}p^5q^2 + \binom{7}{6}p^6q + \binom{7}{7}p^7 =$ $35(1/128) + 21(1/128) + 7(1/128) + 1(1/128) = 64/128 = 1/2$.

(The second method is probably the easier solution to the problem.)

Exercises

1. Find the probability of getting exactly 4 heads in 6 tosses of a fair coin.
2. Find the probability of getting *(a)* a three or five in exactly 2 out of 6 rolls of a fair die; *(b)* exactly 2 even numbers in 5 rolls of a fair die; *(c)* 6 fives in 12 rolls of a die.
3. A card is chosen randomly and returned. This is repeated several times. Find the probability of getting *(a)* 4 hearts in 5 such trials; *(b)* 1 ace in 5 trials; *(c)* at least 1 ace in 5 trials; *(d)* 4 picture cards in 6 trials.
4. A family has 4 children. Find the probability that *(a)* 3 are girls; *(b)* none are girls; *(c)* at least 2 are girls.
5. A woman had 14 children. Find the probability that she had *(a)* at least 12 boys; *(b)* at least 12 girls.
6. In a certain state, 60% of all marriages end in divorce. Find the probability that at least 15 of the 18 marriages announced in today's newspaper of that state end in divorce.
7. In a certain school, 25% of all the students take an extra year to graduate. Find the probability that 10 of Ms. Kress's 25 students do not take that extra year.

PROB

Summary of Combinatorics Problem-Solving Techniques

1. Draw a tree diagram.
2. For any situation that can be interpreted as an AND process, use the fundamental counting principle and multiply to find the total number of possibilities.
3. Look back to check for special cases.
4. Read the problem very carefully until you understand exactly what is being asked, and be sure to clear up any ambiguity before attempting to solve the problem.
5. Do the complementary problem and subtract this answer from the total number of outcomes in general.
6. For any situation that can be interpreted as an OR process, handle each possible case separately and then add to find the total number of possi-

bilities. In other words, if two events are mutually exclusive, and the first event can occur in x ways and the second can occur in y ways, then the total number of ways that the first event or the second event can occur is their sum, $x + y$.

7. Each case of an OR process can often be interpreted as an AND process.
8. For a problem involving an AND process and two independent events A and B, $P(A \wedge B) = P(A) \times P(B)$.
9. For a problem involving an OR process and two mutually exclusive events A and B, $P(A \vee B) = P(A) + P(B)$.
10. If a problem involves n repeated Bernoulli trials, use binomial probabilities.
11. Use a calculator to find most binomial probabilities.

References

Blakeslee, D., & Chinn, W. (1971). *Introductory statistics and probability: A basis for decision making.* Boston: Houghton Mifflin.

Coppins, R., & Umberger, P. (1986). *Applied finite mathematics.* Reading, MA: Addison-Wesley.

Elgarten, G., Posamentier, A., & Moresh, S. (1986). *Using computers in mathematics.* Menlo Park, CA: Addison-Wesley.

Garfunkel, S. (1988). *For all practical purposes: Introduction to contemporary mathematics.* New York: W. H. Freeman.

Goldberg, S. (1960). *Probability: An introduction.* Englewood Cliffs, NJ: Prentice-Hall.

Kemeny, J., Snell, J., & Thompson, G. (1966). *Introduction to finite mathematics* (2nd ed). Englewood Cliffs, NJ: Prentice-Hall.

Kenny, M., & Hirsch, C. (1991). *Discrete mathematics across the curriculum, K-12: 1991 yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Mosteller, F., Rourke, R., & Thomas, G., Jr. (1961). *Probability: A first course.* Reading, MA: Addison-Wesley.

Niven, I. (1965). *Mathematics of choice, or how to count without counting.* Washington, DC: New Mathematics Library, Mathematical Association of America.

Rabinowitz, S. (1992). *Index to mathematical problems 1980-1984.* Westford, MA: MathPro Press.

Soifer, A. (1987). *Mathematics as problem solving.* Colorado Springs: Center for Excellence in Mathematical Education.

Székely, G. (1986). *Paradoxes in probability theory and mathematical statistics.* Budapest: Akadémiai Kiadó.

12

Problem Solving by the Use of Functions

WOLFGANG SCHULZ

Mathematical problems sometimes can be solved successfully by connecting them with the concept of functions. One has to find whether a particular matter can be modeled with the help of a function. If the problem to be solved can be described by the characteristics of a function, then one can try to make use of the knowledge of functions in the solution of the problem.

Some simple examples will illustrate how it is possible to apply the use of functions successfully to problem solving. The stage of modeling will turn out to be especially important and largely decisive for the success of a solution.

For solutions that include the application of functions, it is often important to find first the type of function describing the particular matter or even to find an equation for the function. One can sometimes easily find a solution based on this. Hence, we will discuss some especially elementary—and thus interesting for schools—possibilities for recognizing, even without using terms and methods of calculus, what type of function or what particular function is suitable for modeling, with the goal of finding solutions to the problem.

Some of the following examples, of course, can be treated successfully without deliberately referring to the concept of functions.

Problem 1

ALG

A candle is burning. It originally had a length of 15 cm. After 20 minutes it is only 13 cm long. How long will the candle be burning altogether? How long will the candle remain burning (after these 20 minutes)?

This matter can be modeled with the help of a function. The burning time t can be related to the length of the candle $f(t)$. Assuming that the candle will *always* become 2 cm shorter each 20 minutes, we can presume a *linear* function. This idealizing is very important here; it now makes a very elementary solution possible.

The question for the burning time of the candle is answered if the zero of the function is determined. This problem can be solved through different methods.

⑥ *Making a Picture*

As the graphic representation of a linear function gives points on a straight line, and a straight line is determined by two points, two pairs of values are sufficient for the graphic solution. The pair (0, 15) contains the information that the candle was 15 cm long at the beginning (burning time 0). As we assume that the candle becomes 2 cm shorter each 20 minutes, we get, for instance, the pair (60, 9), because the candle will become 6 cm shorter after 1 hour.

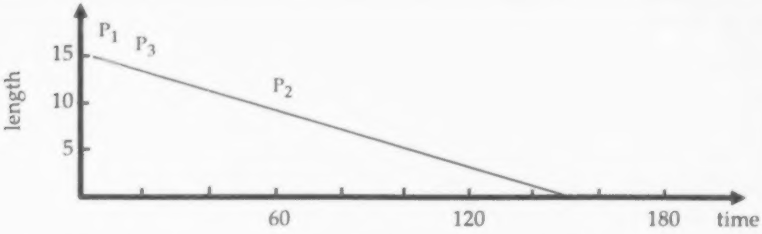


Figure 12.1

If we plot the points P_1 (0, 15) and P_2 (60, 9) into a system of coordinates and draw a straight line through these points, we can read the zero. Hence, it follows that $f(150) = 0$. So the candle is burning altogether 2.5 hours.

The question of how long the candle will remain burning can be answered graphically as well. Candle length 13 cm corresponds to the burning time 20 minutes (see point P_3). Hence the candle will remain burning another 130 minutes.

It is remarkable in this example that it is not necessary to know an equation for the linear function modeling the matter.

Making a Table

⑥

The concept of functions can also help without a graphic representation. The problem can be solved with a table of values. One considers the starting length of the candle (15 cm) and the regular shortening (2 cm every 20 minutes).

Table 12.1

Time t (minutes)	0	20	40	60	80	100	120	140
Length $f(t)$ (cm)	15	13	11	9	7	5	3	1

Proceeding on the assumption that the candle becomes 1 cm shorter every 10 minutes, we get the pair (150, 0). With the help of this pair and the pair (20, 13), one can answer the questions.

Here, too, one does not need an equation for the function. This method is often applied by students without their really recognizing that they work essentially with the function concept.

Forming and Solving an Equation

Since there is an equation $f(t) = at + b$ for every linear function f, where a and b are fixed real numbers, the problem can also be solved with the help of an equation. First one has to determine the numbers a and b. These result from known data.

From $f(0) = 15$ and $f(0) = a \times 0 + b$ follows $b = 15$.

From $f(20) = 13$ and $f(20) = a \times 20 + b$ follows $13 = a \times 20 + 15$, so $-2 = a \times 20$, from which follows immediately $a = -2/20 = -1/10$. Hence, $f(t) = (-1/10)t + 15$. The time t is given in minutes and the length of the candle $f(t)$ in cm.

The slope a can be determined in a different way, too. As the candle becomes shorter, a has to be negative. Since 20 minutes are necessary for a shortening of 2 cm, $a = -2/20 = -1/10$.

The question, How long is the candle burning altogether? means, in the language of functions, For what value of t is $f(t) = 0$? We therefore get the equation

$$-\frac{1}{10}t + 15 = 0$$

$$t = 150$$

The candle is burning altogether for 2.5 hours. How long will the candle remain burning? We only have to subtract the elapsed time (20 minutes) from the total burning time (150 minutes).

All of these methods are mathematically equivalent and prove that there can be more than one promising way to solve a problem. It is an important aim of mathematics classes to make the students aware of this frequently found variety of possibilities.

Exercises

1. In problem 1, we required that the candle shorten proportionately in equal periods of time. We can assume this for candles with constant cross-sections. Make a system of coordinates for the following shapes of candles and plot the lengths of the candles depending on the burning time.

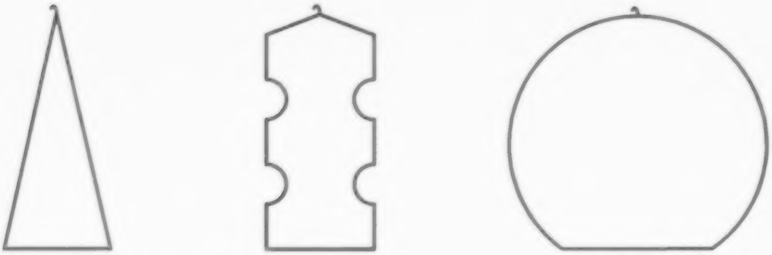


Figure 12.2

2. Which of the following graphic representations can describe the length of a candle depending on the burning time?

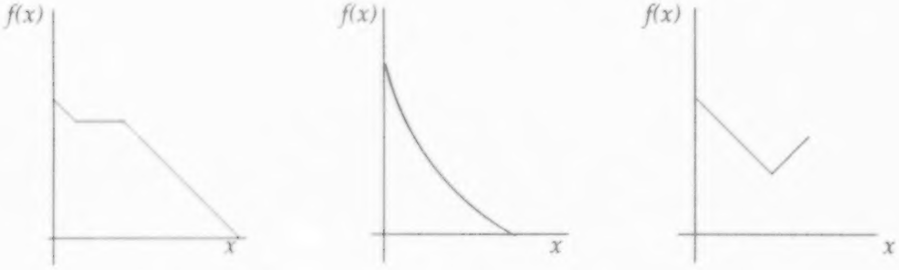


Figure 12.3

3. A swimming pool with a capacity of 150,000 liters is to be filled with water. A pump delivers 1,000 liters per minute. How long will the pump have to continue to work if there are already 30,000 liters of water in the pool? How long will it take to fill the pool to capacity?

A function is sometimes not sufficient to describe a matter. This is the case in many problems when two different motions play a part. If one succeeds in modeling the matter with two suitable functions, then perhaps one can solve the problem by the application of knowledge about functions.

Problem 2

> A military convoy leaves the barracks at 9:00 a.m. for a training exercise and goes at a speed of 45 km/h. At 9:30 a.m., a motorcyclist follows the convoy at a speed of 50 km/h, because the commander forgot to take his orders. When will the motorcyclist overtake the convoy?

ALG

The distance of the convoy from the barracks at the time t (in hours) can be described with the help of the function f. The distance (in km) of the convoy from the barracks at the time t will be stated by $f(t)$. The distance of the motorcyclist from the barracks at the time t (relative to the departure time of the convoy) can be described with the help of the function g. So $g(t)$ will state the distance of the motorcycle from the barracks at the time t. The problem is solved when it is known at which time t both functions have the same values. If the given speeds are constant speeds—which is sensible if one interprets the speeds as average speeds—the functions of f and g are linear. So, again, the situation is simplified by idealizing. Problem 2 can be solved using the same variants of the solution used in problem 1.

Making a Picture

⑥

The graphs of the functions of f and g can be plotted with two pairs of values each. The pairs of values result from the given speeds and the starting times. The motorcycle overtakes the convoy if the graphs of the functions f and g intersect each other.

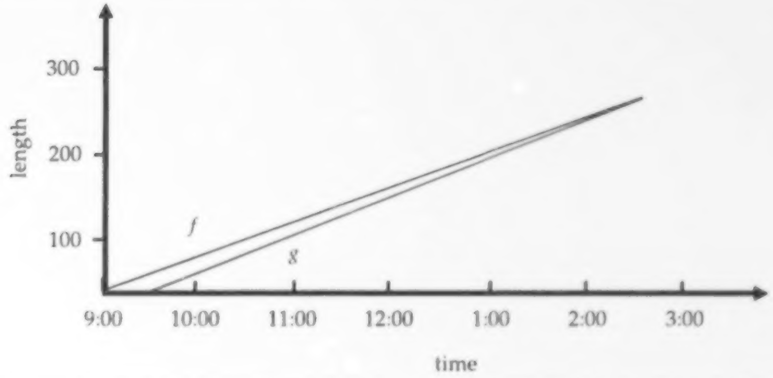


Figure 12.4

Result: The commander can be given his orders around 2:00 p.m.

The above graph makes one general shortcoming of graphic solutions clear: The exactness of the result is not known. Furthermore, one cannot read the point of intersection well if one chooses a clumsy scale, as is the case here. (It is valuable for mathematics teachers to show the limitations of a method of solution, as is the case here.) One should not get too impressed by the often only apparent exactness of calculated results either. In our example, it can be traced back to the rough simplification of the constant speeds.

⑥ *Making a Table*

This problem also can be solved with the help of a table of values. For this purpose, the distances of the convoy and of the motorcycle from the barracks are determined at different times—for example (and for convenience), at every full hour. One has to consider that the motorcycle is traveling 25 km in half an hour (traveling at a constant speed).

Table 12.2

Time t (hours)	9:00	10:00	11:00	12:00	1:00	2:00
Distance $f(t)$ (km)	0	45	90	135	180	225
Distance $g(t)$ (km)		25	75	125	175	225

The result can be read immediately in the last column. If this is not the case, one can try calculating distances for shorter periods of time than one hour. According to this calculation, the meeting time is exactly at 2:00 p.m. But because of the idealization that was made, it is sensible here as well to say that the convoy and the motorcyclist meet at *about* 2:00.

Forming and Solving a Pair of Equations ③

Those for whom the graphic solution is not exact enough and the method of trial and error is too clumsy can find an equation for both functions and then solve the system of equations.

For the convoy, $f(t) = 45t$ (t in hours, $f(t)$ in km) applies.

For the motorcyclist, $g(t) = 50(t - 0.5)$ applies (in consideration of the later start time).

The meeting point is reached when both parties have covered the same distance:

$$\begin{aligned} f(t) &= g(t) \\ 45t &= 50(t - 0.5) \\ t &= 5 \end{aligned}$$

Result: The motorcycle overtakes the convoy 5 hours after the departure of the convoy, that is, at 2:00 p.m.

The exactness of this result should, again, not be overrated. The result is only true for the supposed constant speeds.

Exercises

4. An express train leaves Mountain Valley for Beach Town at 8:00 a.m. It goes the distance of 300 km at a speed of 120 km/h. At 8:45 a.m., a freight train leaves Beach Town for Mountain Valley. It goes at a speed of 80 km/h. When do both trains meet? Do they meet nearer to Mountain Valley or to Beach Town?
5. Mr. Smith is a sales representative. His earnings consist of a fixed salary of $1,000 and a sales commission of 5%. He gets a new offer: a fixed salary of $800 and a sales commission of 7%. For what amount of sales is the new offer more favorable to Mr. Smith?

The first step in solving a problem is often to show patterns in the relationship between the quantities involved. Such patterns often occur in functional interdependencies, provided one has enough data. In the following, two values x and y will be involved, and k pairs of values will be known. We consider only numerical values. One gets the following:

Table 12.3

x	x_1	x_2	x_3	...	x_k
y	y_1	y_2	y_3	...	y_k

The following propositions make the recognition of patterns possible:

⑧ *Proposition 1. The dependence of* y *on* x *is a direct proportion if and only if the quotient* y_n/x_n $(x_n \neq 0)$ *is always constant and equals* a, *and if* $x_n = 0$ *then* $y_n = 0$.

If this is the case, then $y = ax$.

⑧ *Proposition 2. The dependence of* y *on* x *is an inverse proportion if and only if the product* $x_n y_n$ *is always constant and equals* a $(a \neq 0)$.

If this is the case, then $y = a(1/x)$.

⑧ *Proposition 3. The dependence of* y *on* x *is a linear function if and only if the quotient* $(y_i - y_j)/(x_i - x_j)$ *is always constant (for* $x_i \neq x_j$*) and equals* a.

If this is the case, then $y = ax + b$, where $b = y_j - ax_j$ for any j between 1 and k.

⑧ *Proposition 4. The dependence of* y *on* x *is a quadratic function if and only if there are numbers* a *and* b *so that for all* i ≠ j *holds true:* $(y_i - y_j)/(x_i - x_j) = a(x_i + x_j) + b$.

If this is the case, then $y = ax^2 + bx + c$, where $c = y_i - ax_i^2 - bx_i$ for any i between 1 and k.

The propositions 1 through 4 are so strong that one can use them to decide if there is such a pattern or not.

Problem 3

ALG ② In an investigation of some facts, the following table of values was obtained:

Table 12.4

x	0	2	3	5	8
y	4	6	13	39	108

What patterns can you find here?

We are dealing with a function since there is exactly one value of y for each value of x. It cannot be a direct proportion, because, according to proposition 1, the pair (0, 4) must not occur. It cannot be an inverse proportion, because, according to proposition 2, it would have to be true, for instance, that $x_5 y_5 = x_2 y_2$; (i.e., $8 \times 108 \neq 2 \times 6$). It cannot be a linear function, because, according to proposition 3, the quotients $(6-4)/(2-0) = 1$ and $(13-6)/(3-2) = 7$ would have to be equal; but they are not. If one continues to form further quotients such as $(108-13)/(8-3) = 19$ and $(39-13)/(5-3) = 13$, one can check with proposition 4 if there is a linear relation for the following table:

Table 12.5

$x_i + x_j$	2	5	11	8
$(y_i - y_j)/(x_i - x_j)$	1	7	19	13

If one forms the quotients according to proposition 3—such as $(7-1)/(5-2)$, $(13-7)/(8-5)$, or $(19-1)/(11-2)$—then we always get 2 as the result. So one can suppose that there is a quadratic relation with $a = 2$. One can now insert two known pairs—(x_i, y_i), (x_j, y_j)—and $a = 2$ into the equation of proposition 4 to get a value of b. This way one gets, for instance, with (3, 13) and (5, 39), $13 = 2 \times 8 + b$. Because $y_1 - 2x_1^2 + 3x_1 = 4 - 2 \times 0 + 3 \times 0 = 4$, it follows that $c = 4$. So a quadratic relation results: $y = 2x^2 - 3x + 4$. By inserting the numbers x_1, x_2, x_3, x_4, and x_5, we can make sure that this pattern is really true for the whole table of values, which is necessary, as the numbers a and b were not determined generally but resulted from examples.

Exercises

6. Test the following tables of values for patterns:

Table 12.6

a.

x	−1	3	5	10
y	−5	−3	−7	−17

b.

x	0	2	6	7
y	3	−1	−9	−10

c.

x	1	2	4	5	6	8
y	1.5	2.2	5.1	7.6	11.4	25.6

Problem 3 could also be solved with the help of a system of equations. Let us assume that there is a quadratic equation, $y = ax^2 + bx + c$, then the following must be true:

③ $y_1 = ax_1^2 + bx_1 + c$, so $4 = a \times 0^2 + b \times 0 + c$, so $c = 4$.
$y_2 = ax_2^2 + bx_2 + c$, so $6 = a \times 2^2 + b \times 2 + c$, so $6 = 4a + 2b + 4$.
$y_3 = ax_3^2 + bx_3 + c$, so $13 = a \times 3^2 + b \times 3 + c$, so $13 = 9a + 3b + 4$.

This system of equations has the solutions $a = 2$, $b = -3$, and $c = 4$, from which follows $y = 2x^2 - 3x + 4$. One should check, then, that the other pairs of values fit into the equation as well. This check is necessary. If one supposed, for instance, in exercise 6b a linear function, then from the first two pairs it would follow that $a = -2$ and $b = 3$. Hence, the relation could be $y = -2x + 3$. This relation is also true for the pair (6, 9) but not for the pair (7, -10).

Exercises

7. Search for patterns in each of the following tables of values:

Table 12.7

a.

x	−5	−3	1	2	5
y	0.5	1.5	3.5	4	5.5

b.

x	−3	−1	1	2	4
y	42	2	2	17	77

c.

x	−5	−3	1	2	5
y	−3.5	2.5	5	3.5	2.5

If a table of values of a function is given, one can sometimes also trace a pattern with the strategy of "making a picture." If one chooses a system of coordinates with equidistant scaling, one transfers the pairs of values as points into a system of coordinates; then one can decide:

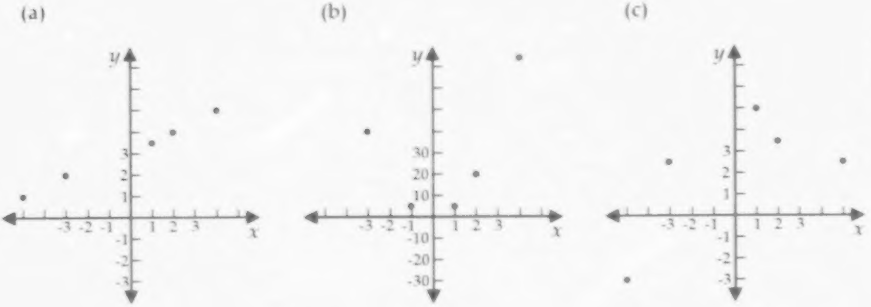
1. There is a direct proportion if and only if the points are all on the same straight line, which runs through the origin of the coordinates.
2. There is a linear function if and only if the points are all on the same straight line.

If one confirms in this way that the table of values is based on such a function, then one can, if necessary, determine the parameters for an equation of functions from the pairs of values.

Other functions—such as inverse proportions, quadratic functions, and exponential and logarithmic functions—cannot be recognized as easily in this way, because their graphs (in a system of coordinates with equidistant scaling) result in points on special curves.

Problem 4

The tables of values for exercise 7 are shown in the graphs below. Is it possible to establish certain types of functions from these graphs?



ALG
⑥
②

Figure 12.5

One can see in *(a)* that all points are on the same straight line. Hence, there is a linear function.

One can see that *(b)* does not resemble a linear function. It could be a quadratic function.

One can see that *(c)* shows none of the functions described in the propositions 1 through 4, because the points are neither on a straight line nor on a curved line that forms a parabola or hyperbola.

Like exponential and logarithmic functions, power functions have such strong characteristics that one can recognize them. Those readers who want to deal more thoroughly with this interesting mathematical problem should refer to J. Aczel's *Lectures on Functional Equations and Their Applications* (New York: Academic Press, 1966). Apart from functional equations, demands on continuity play an important part here.

The following propositions renounce continuity and work with strict monotonicity, which seems more suitable for school, as one can work more easily with the characteristics of monotonicity than with the mathematically demanding term of continuity.

In the following, R denotes the set of real numbers and R_+ the set of positive real numbers.

⑧ *Proposition 5. A function $f: R_+ \rightarrow R$ is a power function with $f(x) = x^k$, $k \in R$ if and only if it is strictly monotonic and $f(x_1x_2) = f(x_1) \times f(x_2)$ for all $x_1, x_2 \in R_+$.*

If this is the case, *k* can be determined by solving the equation $f(x_1) = x_1^{\,k}$ for an $x_1 \neq 1$.

⑧ *Proposition 6. A function $f: R_+ \rightarrow R$ is an exponential function with $f(x) = a^x$, $a \in R$ if and only if it is strictly monotonic and $f(x_1 + x_2) = f(x_1) \times f(x_2)$ for all $x_1, x_2 \in R$.*

If this is the case, the number *a* results from the functional value *f*(1) or from the solution of an equation $f(x_1) = a^{x_1}$ for an $x_1 \neq 0$.

⑧ *Proposition 7. A function $f: R_+ \rightarrow R$ is a logarithmic function with $f(x) = \log a^{x}{}_{1}$ $a \in R$ if and only if it is strictly monotonic and $f(x_1x_2) = f(x_1) + f(x_2)$* for all $x_1, x_2 \in R_+$.

If this is the case, *k* can be determined by solving the equation $f(x_1) = x_1^{\,k}$ for $x_1 \neq 1$.

Problem 5

ALG

② A bacterial culture is growing. Observations for several days have resulted in the following table:

Table 12.8

Time *t* (days)	0	1	2	3	4	5	6	7
Mass *f*(*t*) (mg)	1.0	1.5	2.2	3.4	5.1.	7.6	11.4	25.6

What pattern does the process of growing show? Is it possible to make predictions for further development? (See also exercise 6c.)

From the table of values it follows that the function *f* is strictly monotonically growing. One can check quickly that the conditions of propositions 1 through 4 are not met. Nor are the conditions of propositions 5 met, because there is a clear difference if one calculates, for instance, $f(2 \times 3) = f(6) = 11.4$ and compares this number with $f(2) \times f(3) = 2.2 \times 3.4 = 7.48$. Hence, it cannot be a power function.

If one calculates $f(2 \times 3) = f(6) = 11.4$ and compares this number with $f(2) + f(3) = 2.2 + 3.4 = 5.6$, by proposition 7, we can conclude that it is not a logarithmic function either.

Now, the following numbers are compared:

$f(1 + 1) = f(2) = 2.2$ and $f(1) \times f(1) = 1.5 \times 1.5 = 2.25$
$f(1 + 2) = f(3) = 3.4$ and $f(1) \times f(2) = 1.5 \times 2.2 = 3.3$
$f(2 + 2) = f(4) = 5.1$ and $f(2) \times f(2) = 2.2 \times 2.2 = 4.84$
$f(2 + 3) = f(5) = 7.6$ and $f(2) \times f(3) = 2.2 \times 3.4 = 7.48$
$f(2 + 4) = f(6) = 11.4$ and $f(2) \times f(4) = 2.2 \times 5.1 = 11.22$

The numbers are different; hence, according to proposition 6, it is not an exponential function either. But they do not differ very much from each other. Since it was worked with measured values, one has to consider inaccuracies caused by errors in measurement. It is therefore sensible to try, if possible, to describe this process adequately by an exponential function; $f(1) = 1.5$ presents itself as a suitable base a of the exponential function; hence, $f(t) = 1.5^t$. If we compare now the measured values $f(t)$ with the functional values 1.5^t (rounding to two decimal places), we get:

Table 12.9

t	0	1	2	3	4	5	6	8
$f(t)$	1.0	1.5	2.2	3.4	5.1	7.6	11.4	25.6
1.5^t	1.0	1.5	2.25	3.38	5.06	7.59	11.39	25.63

⑥

One can say that this function is well suited as an approximation for the process of growing. This function can now also be used to obtain further knowledge about the growth process. It follows, for instance, that $1.5^7 \approx 17.1$, which is an approximation for the mass of bacteria after 7 days. One should be careful of long-term predictions. After 20 days, there results a mass of bacteria of 1.5^{20} mg = 3,325 mg. This forecast comes true only if the conditions for the process of growing do not change—for instance, there has to be enough space and food for the bacteria to grow.

Exercises

8. The following table shows the number of people on Earth for certain years.

Table 12.10

Year	1900	1950	1960	1980	1985
Number (in millions)	1,608	2,483	3,014	4,453	4,837

Is it possible to describe the growth of the population with a function? Are predictions for further growth possible? You should compare calculated values with up-to-date figures as well.

9. Test the following tables of values for patterns:

Table 12.11

a.

x	1	1.5	2	2.5	3	3.5	4
y	0	0.4	0.7	0.9	1.1	1.3	1.4

b.

x	1	2	3	4	6	8	9
y	1.00	1.26	1.44	1.59	1.82	2.00	2.08

c.

y	–2.5	–1	0	1	2	3	6
y	15.5	3	1	0.333	0.11	0.04	0.001

10. Calculate the mass of the bacteria in problem 5 after 12 hours, after 9 days, and one day before the observation started, with the help of

    (a) the functional equation $f(x_1 + x_2) = f(x_1) \times f(x_2)$ and already known values, and
    (b) the functional equation $f(t) = 1.5^t$.

(Note: If one works with systems of coordinates with logarithmic scaling of one or both axes, one can graphically recognize additional functions, because the graphs of the exponential and logarithmic functions are then straight lines.)

Propositions 1 through 7 express an idealistic state, which will hardly occur in reality (see problem 5). However, unpreventable errors in measurement, accidents, and impairments provide numbers that do not comply with the above propositions. Nevertheless, one can make some sense of these propositions.

If one is working with proposition 3, and it turns out that the quotients $(y_1 - y_j)/(x_i - x_j)$ differ only slightly from each other, one can determine a number *a* with the average of *all* possible quotients. The average, then, of *all* numbers $y_i - ax_i$ gives the number *b*. This way, a linear function *f* with $f(x) = ax + b$ comes into being and has to be checked for suitability for the description of the matter.

Problem 6

Immediately after birth the length and weight of babies is determined. Is it possible to say "the longer the heavier"? Is there a linear function describing the relation between the length and the weight?

The following data result from measurements on newborn babies:

STAT

GEN

②

Table 12.12

Length *l* (cm)	51	53	53	54	50	52	48	52	49	49
Weight *w* (g)	3,190	3,450	3,840	3,770	3,260	3,020	2,620	3,930	3,000	2,920

From this table it follows immediately that there is no functional relation between length and weight because there are different weights for equal lengths. Here is a stochastic relation. But also in such a case one tries to find an approximately linear correlation.

This can happen graphically by representing the pairs as points in a system of coordinates and then plotting a straight line for which the distances between the points and the straight line are as small as possible. This is often done quite well by eye. By way of calculation, the method of least squares has been known since Carl Friedrich Gauss. A foundation of this method in school is difficult. One can try it in an elementary way, but this may be quite clumsy; and if one uses calculus, one needs the theory of extreme values of several variables. If one organizes the data that way—that for equal lengths the average weight is used—one gets the following table:

Table 12.13

Length *l* (cm)	48	49	50	51	52	53	54
Weight *w* (g)	2,620	2,960	3,260	3,190	3,475	3,645	3,770

These data yield a function. We can recognize a tendency for children who are taller to be heavier as well.

If one now forms the quotient of proposition 3 in the above-mentioned way and then calculates the numbers a and b, one will get the linear function $w = 174l - 5{,}600$ (after rounding). If one inserts the measured values into this equation and considers the deviations, one can judge whether or not the linear function found gives a sufficient description of the matter. It is obvious that the consideration of a larger amount of data will give a better result. If the investigation is based on a large amount of such data, one gets a linear function that describes the correlation between length and weight of newborn babies on average.

Exercise

11. With the help of the following data, one can examine whether there are noticeable differences between newborn boys and girls. Which babies' measurements differ significantly from the average?

Table 12.14

Girls									
Length l (cm)	54	51	54	52	48	52	49	53	50
Weight w (g)	3,620	2,800	3,590	2,800	2,210	4,170	2,980	3,770	2,950

Boys									
Length l (cm)	51	53	53	56	54	52	51	50	47
Weight w (g)	2,400	3,900	4,350	3,750	3,920	3,930	3,240	3,870	2,610

From the techniques presented in this chapter, one should begin to get a genuine appreciation for the role that functions can play in solving problems that might normally be solved in other ways. At the elementary levels, the techniques should be more familiar, as the beginning of the chapter showed. Yet these more familiar, or common, uses have more sophisticated applications. The reader is encouraged to pursue further applications according to individual interests.

13

Symmetry Saves the Solution

DAVID SINGMASTER

Symmetry is a powerful tool in solving mathematical problems. Although symmetric conditions need not give symmetric answers (see Stewart, 1992, for examples), there are many cases where exploitation of the symmetry leads to solutions that otherwise are much harder or even impossible to find. In addition, the solution and the steps leading to it are often made more intelligible by the light of the symmetry. In this essay, I will describe some recreational problems where I have found symmetry useful and even essential.

The Two Towers

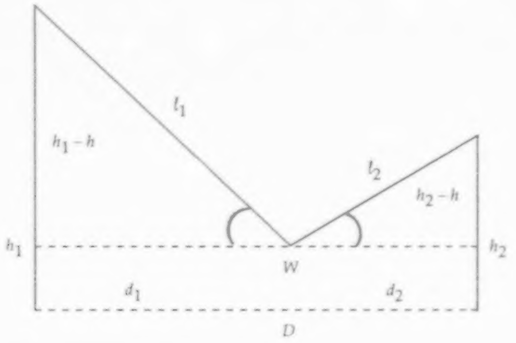
GEN

GEOM

Several medieval arithmetic/algebra texts give a problem of locating a fountain between two towers so that pigeons of equal speeds can get from the tops of the towers to the fountain in equal times. The solution to it is straightforward—indeed, it reduces to a linear problem. However in the *Trattato d' Aritmetica* (c. 1370), attributed to Paolo dell'Abbaco, there are problems in which a rope is strung between the two tower tops and a weight is hung from a ring on the rope, which is just long enough for the weight to reach the ground (dell'Abbaco, 1964, pp. 129-133). Solving this is trickier than the preceding version. To make it even trickier, suppose the rope is not long enough—then how high from the ground does the weight hang? Consider Figure 13.1.

⑱

⑥



ALG

Figure 13.1

We are given h_1, h_2, $D = d_1 + d_2$ and $L = l_1 + l_2$ and the problem is to find h. Since the weight can slide on the rope, the two marked angles at W are equal and ⑬ we have the following five equations in the five unknowns l_1, l_2, d_1, d_2, and h.

$$l_1 + l_2 = L$$

$$d_1 + d_2 = D$$

$$(h_1 - h)^2 + d_1^2 = l_1^2$$

$$(h_2 - h)^2 + d_2^2 = l_2^2$$

$$\frac{l_1}{d_1} = \frac{l_2}{d_2}$$

③ It took me several tries to solve this, but I have now found a reasonably direct solution. Using $a/b = c/d$ if and only if $a/b = (a + c)/(b + d)$, we deduce $L/D = l_1/d_1 = l_2/d_2$.

Dividing the third of the given equations by d_1^2, and using the previous sentence, we have

$$\left(\frac{h_1 - h}{d_1}\right)^2 + 1 = \left(\frac{l_1}{d_1}\right)^2 = \left(\frac{L}{D}\right)^2.$$

Symmetrically, we have

$$\left(\frac{h_2-h}{d_2}\right)^2+1=\left(\frac{L}{D}\right)^2.$$

Hence

$$\frac{h_1-h}{d_1}=\frac{h_2-h}{d_2}=\sqrt{\frac{L^2-D^2}{D^2}},$$

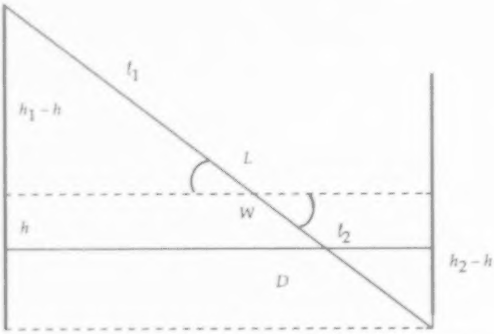
and so

$$\frac{h_1-h+h_2-h}{d_1+d_2}=\sqrt{\frac{L^2-D^2}{D^2}};$$

so

$$(h_1+h_2-2h)^2=L^2-D^2.$$

The simplicity of this result naturally makes one try to find a simple derivation, and symmetry does it! Reflect one tower in the horizontal through *W* to get Figure 13.2. ③



① ② ⑥

Figure 13.2

The two parts of the rope now form a straight line that is the hypotenuse of a right triangle with sides *D* and $(h_1-h)+(h_2-h)$, so we immediately have $L^2=(h_1+h_2-2h)^2+D^2$.

Some Historical Notes

The earliest known example of the simpler version with the pigeons and the fountain is in Fibonacci (1857, pp. 331-332, 398-399). Vogel's (1970-1977) article on Fibonacci says the problem is of Indian origin, but I have not located any such versions.

Dell' Abbaco only considers the case $h = 0$ and proceeds by computing $H = h_1 + h_2$, $h_1L/H\ (= l_1)$ and then $(h_1\ L/H)^2 - h_1^2 = d_1^2$. He gives no geometric justification, so the problem must have been at least somewhat known. However, I have found no other occurrences of the problem until 1778 (Ozanam, 1778, p. 11), and I have found only one other occurrence, in a mid-19th-century algebra text. Before discovering the problem in Ozanam, I posed the problem in *Crux Mathematicorum* (Singmaster, 1992) and then discovered that Ozanam's solution is the same as mine.

Square Hunting

Klamkin, Breault, and Schwartz (1959-1960) posed and solved the following problem:

GEN

> An explorer travels on the surface of the earth, assumed to be a perfect sphere, in the manner to be described. First, he travels 100 miles due north. He then travels 100 miles due east. Next he travels 100 miles due south. Finally, he travels 100 miles due west, ending at the point from which he started. Determine all the possible points from which he could have started.

ALG

GEOM

TRIG

The first solver gave only the obvious solution: 50 miles south of the equator. The later solver found the general solution but could not find the answers in closed form.

In 1983-1984, the problem and its solution appeared in the *Journal of Recreational Mathematics* (Kakinuma, Barwell, and Collins), and Brian Barwell found a better solution but did not quite carry it to its conclusion.

The only other version of the problem that I have seen is Perelman (1985, pp. 18-19, 25-26), but I do not know when this first appeared. Perelman has a dirigible (later a helicopter) traveling 500 km north, east, south, and west from Leningrad—where does it land?

About 1986, the problem recurred to me, though I could not remember where it had come from. I found the following closed-form solution by use of a standard symmetry before I found the literature.

⑥

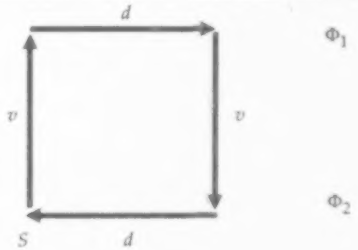


Figure 13.3

We generalize somewhat, as shown in Figure 13.3. We start at S and go v to the north, then d to the east, v to the south and d to the west, and return to our starting point. Let the colatitude (= polar angle) of the eastward trip be Φ_1, and that of the westward trip be Φ_2. ⑱

Let R be the radius of the earth. Since arc length is radius times angle, we have

$$\Phi_2 - \Phi_1 = v/R. \qquad [1]$$

Traveling at constant colatitude Φ is to travel around a circle of radius $R\sin\Phi$, hence the longitude covered by traveling distance d at colatitude Φ is $d/R\sin\Phi$, so we want

$$\frac{d}{R\sin\Phi_1} \equiv \frac{d}{R\sin\Phi_2} \ (\text{mod}2\,\pi) \qquad [2]$$

or

$$\frac{d}{R\sin\Phi_1} = \frac{d}{R\sin\Phi_2} + 2\pi k \qquad [2']$$

for some integer k, which is the number of times the explorer circles the earth. The solution of [1] and [2] or [2′] is not obvious and frustrated the 1960 solver. Assuming we are not at the special cases $\Phi_1 = 0$ or $\Phi_2 = \pi$, we can symmetrize the problem by setting ⑰

$$\Phi = \frac{\Phi_1 + \Phi_2}{2} \qquad \beta = \frac{\Phi_2 - \Phi_1}{2} \qquad [3]$$

so

$$\Phi_1 = \Phi - \beta, \quad \Phi_2 = \Phi + \beta\,. \qquad [4]$$

(Barwell [Kakinuma, Barwell, & Collins, 1983-1984] does the same.) Setting $\alpha = d/2R$ and using several trigonometric identities, [2′] becomes

$$\frac{\pi k}{\alpha} = \frac{1}{\sin(\Phi-\beta)} - \frac{1}{\sin(\Phi+\beta)} = \frac{\sin(\Phi+\beta) - \sin(\Phi-\beta)}{\sin(\Phi-\beta)\sin(\Phi+\beta)} \qquad [5]$$

TNS

$$= 2\,\frac{\cos\Phi\,\sin\beta}{\frac{1}{2}\cos2\beta - \frac{1}{2}\cos2\Phi} = 2\,\frac{\cos\Phi\,\sin\beta}{\cos^2\beta - \cos^2\Phi}\,.$$

Assuming again that we are not in the special case $k = 0$, we can continue. Cross-multiplying equation [5] gives a quadratic equation whose solution is

$$\cos\Phi = \frac{-\alpha\sin\beta \pm \sqrt{\alpha^2\sin^2\beta + \pi^2k^2\cos^2\beta}}{\pi k}\,. \qquad [6]$$

(Barwell [Kakinuma, Barwell, & Collins, 1983-1984] got to the quadratic but thought the solution too messy to consider and went to consider approximate solutions.)

This solution leads to a number of further questions and variations of the problem, which I have treated in an unpublished paper. I give one example below. None of this would have been possible without the basic symmetrization step of changing from Φ_1 and $\Phi_2 = \Phi_1 + v/R$ to $\Phi - \beta$ and $\Phi + \beta$. This process goes back to
⑱ the Babylonians, who used it to solve such problems as $x + y = 10$, $xy = 24$ by setting $x = 5 + \beta$, $y = 5 - \beta$, so $xy = 25 - \beta^2 = 24$ and $\beta^2 = 1$. (See p. 86)

This symmetrization has the effect of eliminating the first-order term in a
③ quadratic and is equivalent to completing the square. The square bear problem is a trigonometric problem and hence more complicated, but the symmetrization is even more essential.

As an example of a new problem inspired by the above, consider the following (Singmaster, 1993):

> ⑰ Hiawatha, the mighty hunter, has traveled far in search of game. One morning he gets up, has breakfast, heads north and travels 10 miles forward in a straight line. Seeing nothing, he stops for lunch. After lunch, he heads north and travels 10 miles forward in a straight line and finds himself where he started in the morning. Where on earth is he?

Men Buy a Horse

GEN
ALG

> Two men wish to buy something—usually a horse—but neither has enough money to do so. The first says to the second, "If I had half of your money, I could buy the horse." The second responds, "If I had a third of your money, I could buy the horse."

If the value of the horse is known, this is a determinate problem. But the value is often not given and the problem is then indeterminate, and one wants the smallest integral solution or a general solution. Such problems date back to the *Chiu Chang Suan Ching* (1968, pp. 86-88, problems 10, 12, 13), written c. A.D. 100. The above example occurs there with an unspecified object worth 50. Diophantos, c. A.D. 250 (Heath, 1964, pp. 139-149), gives a somewhat general approach. Historically, these problems are significant as being among the first where negative solutions are treated naturally—for example, by Fibonacci (1857, pp. 228-258, 327-349) in 1202.

When one extends to more than two people, there are two forms of the problem.

> *Type I:* The i^{th} person says, "If I had a_i times what all of you have, then I could buy the horse."
>
> *Type II:* The i^{th} person says to the $i + 1^{\text{st}}$ person (considered cyclically), "If I had a_i times what you have, then I could buy the horse." ⑱

In both types, all known versions have all a_i rational and we will assume this.

The Type I problem has an obvious symmetry, and symmetry will guide us to a solution. Let x_i be the amount the i^{th} person has, and let T be the total amount they have, that is, $T = \Sigma x_i$. Let h be the value of the horse or other object being purchased. Since $T - x_i$ is the amount that all except the i^{th} person have, our equations can be written as:

$$x_1 + a_i(T - x_i) = h\,. \qquad [7]$$

(The use of the symmetric expression T already makes the problem much simpler.) Solving [7] for x_i gives us

$$x_i = \frac{h - a_i T}{1 - a_i}, \qquad [8]$$

and adding these for all i eliminates the individual x_i's, giving us

$$T = h\Sigma\frac{1}{1-a_i} - T\Sigma\frac{a_i}{1-a_i}, \qquad [9]$$

so ⑬

$$T\left(1 + \Sigma\frac{a_i}{1-a_i}\right) = h\left(\Sigma\frac{1}{1-a_i}\right). \qquad [10]$$ ②

One can readily find the smallest integers T and h satisfying [10], but a more satisfactory expression arises from rewriting [8] as

$$x_i = T + \frac{h - T}{1 - a_i}\,. \qquad [11]$$

Adding these for all i gives

$$T = nT + (h - T)\,\Sigma\frac{1}{1-a_i} \qquad [12]$$

or

$$(n-1)T = (T - h)\,\Sigma\frac{1}{1-a_i}, \qquad [13]$$ ②

which shows that T and $T - h$ are a simpler pair of unknowns than T and h. The
⑰ least integers T and $T - h$ satisfying [13] correspond to the least integers T and h satisfying [10]. For example, the *Chiu Chang Suan Ching* example has $n = 2, a_1 = 1/2, a_2 = 1/3$ so [13] becomes $T = (T - h)7/2$ or $2T = 7(T - h)$, and we can take $T = 7$, $T - h = 2$, giving $h = 5, x_1 = 3, x_2 = 4$.

Fibonacci (1857) devotes a chapter of 30 pages to this problem, giving many variations with up to seven persons or five horses, an inconsistent example, and examples with negative solutions. He treats many of the problems by another method in a later chapter. The numbers involved can be quite formidable—in one case $h = 35{,}839{,}901$.

There is a c. 1390 version (*Libro d'abaco*, 1973, p. 141) with an extra mutual friend, which leads to a two-parameter solution space, though only one solution is given. The easiest example with a negative solution appears in a Provençale manuscript of c. 1430 (*Provençale Arithmétique*, 1984, pp. 49-53) and is just our problem for $n = 5$ people, with $(a_i) = (1/2, 1/3, 1/4, 1/5, 1/6)$.

⑯ Unfortunately, integral values for T and h do *not* guarantee integral values for all x_i, but one can always convert a solution to a solution with integral x_i.

Type II problems give the equations:

$$x_i + a_i x_{i+1} = h. \tag{14}$$

Although not as symmetric as type I problems, we can solve by systematic elimination. Combining [14] for $i = 1$ and $i = 2$ gives us

② $$x_1 + a_1 x_2 = h,$$

$$x_2 + a_2 x_3 = h,$$

so

$$x_2 = h - a_2 x_3,$$

and

$$x_1 + a_1(h - a_2 x_3) = h$$

and

$$x_1 - a_1 a_2 x_3 = h - a_1 h.$$

④ When $n = 2$, then $x_3 = x_1$, and we have

⑰ $$x_1(1 - a_1 a_2) = h(1 - a_1).$$

Otherwise, we continue with

$$x_3 + a_3 x_4 = h,$$
$$x_3 = h - a_3 x_4,$$
$$x_1 - a_1 a_2 (h - a_3 x_4) = h - a_1 h,$$

getting ②

$$x_1 + a_1 a_2 a_3 x_4 = h(1 - a_1 + a_1 a_2).$$

When $n = 4$, we have

$$x_1(1 + a_1 a_2 a_3) = h(1 - a_1 + a_1 a_2).$$

Continuing, we see, for n people, ⑱

$$x_1 [1 + (-1)^{n+1} a_1 a_2 \dots a_n] = h [1 - a_1 + a_1 a_2 - \dots + (-1)^n a_1 a_2 \dots a_n]. \qquad [15]$$

Trying to determine x_2, x_3, and so on from [14] and [15] gives a mess, but we can simply note that the problem looks the same wherever we start from so we get

$$x_i [1 + (-1)^{n+1} a_i a_{i+1} \dots a_{i-1}] = h [1 - a_i + a_i a_{i+1} - \dots + (-1)^n a_i a_{i+1} \dots a_{i-1}] \qquad [16]$$

Here we have again used symmetry, in a simpler way than in type I problems but just as usefully.

Men Find a Purse and Similar Problems

Here the ith person says, "If I had the purse, I'd have a_i times what you have." ALG
Again, for more than two people, there are two types of problems, depending on whether "you" means "all of you" or just the next person in sequence (taken ⑬
cyclically). In type I, we get the equation $x_i + p = a_i (T - x_i)$, while type II gives the equation $x_i + p = a_i x_{i+1}$.

Here p is the value of the purse. We see that this is the same problem as men ⑱
buying a horse but with a horse of negative value and with negative multipliers! Algebraically, the solution in this situation requires nothing new. However, there are two ancient problems which are related to this form. The simpler gives the ④
following equations for $n = 3$.

$$x_1 + x_2 = a_3$$
$$x_2 + x_3 = a_1$$
$$x_3 + x_1 = a_2$$

⑱ Diophantos (Heath, 1964, p. 135) gives a general approach to this and examples, with $(a_i) = (20, 30, 40)$ and $(22, 24, 27, 30)$. The general form of this has n equations stating that the sum of all but one of the values is a given value. If we let $T = \Sigma x_i$ as before, then the equations are $T - x_i = a_i$. We proceed symmetrically by adding all these to get $nT - T = \Sigma a_i$; hence $T = \frac{1}{n-1} \Sigma a_i$, from which $x_i = T - a_i$ are readily found.

⑱ The other related problem is the "bloom" of Thyramides (see Heath, 1981, pp. 94-96). This has n + 1 unknowns: $x, x_1, x_2, \ldots, x_n$; and we are given $x + x_i = a_i$ and $x + x_1 + \ldots + x_n = S$.

As before, let $T = \Sigma x_i$, so $x + T = S$. We again proceed symmetrically by adding to get $nx + T = \Sigma a_i$.

Combining this with $x + T = S$ gives us $(n-1)x = \Sigma a_i - S$, so x and all the x_i are determined.

Men and a Vehicle

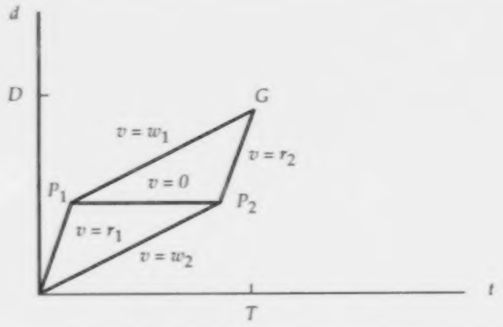
GEN The simplest version of this problem has two men and a bicycle, which cannot carry a passenger. The men wish to get to the next village, a distance D away, as soon as possible. Assuming the men have equal walking speeds and equal riding speeds, it is easy to see that the optimum method is for one to ride halfway and then leave the bicycle for the other to pick up. Thus, each person rides halfway and walks halfway.

The earliest occurrence I've seen of this simple version is in Gaston Boucheny (1939, pp. 77-78), but more complex versions go back to about 1900 (Loyd, 1960, pp. 88, 160-161), and I suspect there must be older examples. [I have since found 1906 versions with two persons and a bicycle and with four persons and a car which can carry two passengers (Laisant, 1906, pp. 127-132).]

Complexity can be introduced in several ways. First, the speeds of the various individuals may differ. Second, we may have more people. Third, the vehicle may carry more people (versions of the problem have a tandem bicycle, a motorcycle with passenger seat, a taxi, a truck or van, or a pony). In all cases, the vehicle cannot carry all the people, so the vehicle must shuttle back and forth—possibly with different speeds when loaded and unloaded. I was led to investigate this type of problem when I saw an elegant graphical solution in Menninger (1961, pp. 100-101). Unfortunately this solution is incorrect, but it led me to the following algebraic solution, and the graph is still the best way to see the equations.

Let us first consider the simplest case of two men and a bicycle. The men walk at rates w_i and ride at rates r_i. We graph position d against time t for both men and the bicycle as in Figure 13.4.

⑬ The first person rides at rate r_1 until he leaves the bicycle, and he then walks at rate w_1. The bicycle stays put until the second person gets to it, and then he rides it the rest of the way. Obviously, the optimum solution has both people arriving at the goal at the same time. Letting $P_1 = (t_1, d_1)$, $P_2 = (t_2, d_2)$ and $G = (T, D)$ in the



⑥

Figure 13.4

diagram, and remembering that distance is velocity × time, we have five unknowns—T, t_1, t_2, d_1, d_2—and the following five equations:

$$d_1 = d_2 \qquad [17]$$

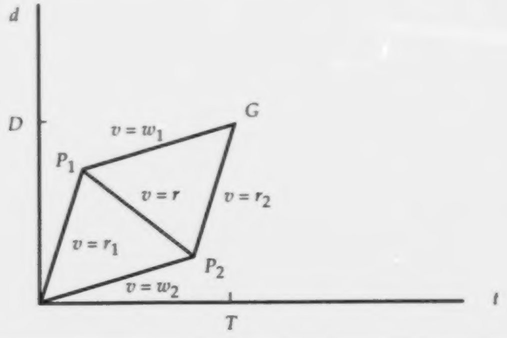
$$d_1 = r_1\, t_1 \qquad [18]$$

$$d_2 = w_2\, t_2 \qquad [19]$$

$$D - d_1 = w_1\,(T - t_1) \qquad [20]$$

$$D - d_2 = r_2\,(T - t_2) \qquad [21]$$

⑱ Although these have some symmetry and can be solved, the process is not elegant, especially if one tries to deal with the more general problems. In particular, if we have a vehicle which can return, then the diagram looks like Figure 13.5, where r is the unloaded return speed of the vehicle.



⑥

Figure 13.5

This changes [17] to

$$d_1 - d_2 = r\,(t_2 - t_1). \qquad [17']$$

③

⑬ After some trial and thought, I realized that the problem becomes much simpler if one uses different variables that are less obvious but turn out to make the problem much more symmetric and hence readily solvable.

We now let t_i be the time the i^{th} person rides. Then each person rides for time t_i and walks for time $T - t_i$ to cover the total distance D, so we have

$$r_i\,t_i + w_i\,(T - t_i) = D. \qquad [22]$$

(If we let $d_i = r_i\,t_i$, then the point P_2 has coordinates $(T - t_2, D - d_2)$, which is not an obvious improvement!) We can solve [22] for t_i as:

$$t_i = \frac{D - w_i T}{r_i - w_i}. \qquad [23]$$

If we have a bicycle (or a pony which does not head back), then it also travels a distance D, giving us

$$\Sigma r_i\,t_i = D. \qquad [24]$$

We insert equations [23] into [24] to get

$$\Sigma \frac{r_i}{r_i - w_i}\,(D - w_i T) = D,$$

and this solves to

$$D\left(\Sigma \frac{r_i}{r_i - w_i} - 1\right) = T\left(\Sigma \frac{r_i\,w_i}{r_i - w_i}\right). \qquad [25]$$

I do not see here any simplification similar to that in equations [11]-[13].

If the bicycle is replaced by an active means of transport such as a motorcycle, then we need to replace [24]. The cycle will move forward $\Sigma r_i\,t_i$, where r_i may vary due to the weight of the passenger. The cycle spends time Σt_i going forward and hence time $T - \Sigma t_i$ going backward. Assuming all its return trips are at the same speed r, it travels $r(T - \Sigma T_i)$ backward. Its net forward travel must be D, which gives us

$$\Sigma r_i\,t_i - r(T - \Sigma t_i) = D. \qquad [24']$$

Combining equations [23] into [24′] gives, after a few steps,

$$D\left(\Sigma \frac{r_i + r}{r_i - w_i} - 1\right) = T\left(\Sigma \frac{(r_i + r)w_i}{r_i - w_i} + r\right). \qquad [26]$$

In the case of constant r_i ($= r$) and constant w_i ($= w$) and a bicycle, [25] reduces to

$$[(n-1)r + w]\, D = nrwT. \qquad [27]$$

Hence, the average speed $V = D/T$ is $V = nrw/[(n-1)\, r + w]$, and we see $1/V = [(n-1)/n]\, [1/w] + (1/n)\, (1/r)$, so V is the appropriately weighted harmonic mean of w and r. This is correct because each person travels D/n by bicycle and $(n-1)\, D/n$ by foot. ② ⑰ ③

If the return speed of the vehicle also has the constant value r of all r_i, then [26] simplifies to $[(2n-1)\, r + w]\, D = [(2n-1)\, rw + r^2]\, T$.

Goodstein (1945) obtains this by a slightly less obvious process. Menninger's problem has three walkers and a motorcyclist who can carry one passenger, $D =$ 5 km, $r_i = r = 30$ km/h, $w_i = 5$ km/h. This leads to $T = 31/66$ hr $= 28.1818$ min. Menninger got 35 min. because he had the cyclist carry the passengers all the way to the next village, forgetting that they could do some walking at that end and destroying the symmetry of the problem!

Concluding Remarks

As I stated at the beginning, symmetry is a powerful and useful tool in problem solving. However, symmetry is a subtle concept and it is not easy to describe how it can be recognized and used. The best way to learn it is by seeing examples, and I have here presented a number of uncommon recreational examples which seem suitable for learning from.

References

Boucheny, G. (1939). *Curiositiés and récréations mathématiques*. Paris: Larousse.

Chiu chang suan ching (Nine chapters on the mathematical art). (1968). (Trans. K. Vogel as *Neun Bücher Arthmetischer Technik*). Braunschweig: Vieweg.

dell'Abbaco, P. (1964). *Trattato d'Aritmetica* (c. 1370). Ed. G. Arrighi Pisa: Domus Galilaeana. (Warren Van Egmond asserts that this manuscript is a 15th-century compilation and doubts that it is due to dell'Abbaco ["New light on Paolo dell'Abbaco," in *Annali dell'Istituto e Museo di Storia della Scienzia di Firenze*, *2*(2), (1977), 3-21].)

Fibonacci, L. (1857). *Scritti di Leonardo Pisano* (Vol. 1, Ed. B. Boncompagni). Rome: B. Boncompagni. (Original publication 1202)

Goodstein, R. L. (1945). Note 1797: Transport problems. *Mathematics Gazette, 29*, 16-17.

Heath, T. L. (1964). *Diophantos of Alexandria*. New York: Dover. (Original publication 1910)

Heath, T. L. (1981). *A history of Greek mathematics* (2 vols.) New York: Dover. (Original publication 1921)

Kakinuma, Y., Barwell, B., & Collins, C. H. (1983-1984). Problem 1212: Variation of the polar bear problem. *Journal of Recreational Mathematics, 16*(3), 226-228.

Klamkin, M., Breault, D. A., & Schwartz, B. L. (1959-1960). Problem 369. *Mathematics Magazine, 33*, 226-228.

Laisant, C. A. (1906). *Initiation mathématique.* Paris: Hachette.

Libro d'abaco. (1973). (Ed. G. Arrighi). Lucca: Cassa di Risparmio. (Original publication c. 1390)

Loyd, S. (1960). *Mathematical puzzles of Sam Loyd* (Vol. 2, Ed. M. Gardner). New York: Dover. (Original publication 1914)

Menninger, K. (1961). *Mathematics in your world.* London: G. Bell. (Original publication 1954)

Ozanam, J. (1778). *Récréations Mathematiques et physiques.* (Revised by J. E. Montucla). (Original publication 1964)

Perelman, Y. I. (1985). *Mathematics can be fun* (3rd ed). Moscow: MIR.

Provençale arithmétique. (1984). (Trans. and annot. by J. Sesiano). *Centaurus, 27*, 26-75.

Singmaster, D. (1992). Problem 1748: The two towers. *Crux Mathematicorum, 18*(5), 140.

Singmaster, D. (1993). The hunting game. *Focus*, No. 3, 77, 98.

Stewart, I. (1992). Curie's mistake. In *Another fine math you've gotten me into* (pp. 145-160). New York: Freeman.

Vogel, K. (1940). Zur geschichte der linearen gleichungen mit mehreren unbekannten. *Deutsche Mathematik, 5*, 217-240.

Vogel, K. (1970-1977). Fibonacci, Leonardo, or Leonardo of Pisa. In *Dictionary of Scientific Biography, 4*, 604-613.

14

An Application of Congruence Transformations in Problem Solving

JAN TROJÁK

There is a whole category of geometric problems consisting of a search for a definite geometric figure, some important points of which are given, while some other points, indispensable for determining the position of the figure, are requested to lie on given geometric objects; in plane geometry, these are usually (though not necessarily) lines or circles or their parts.

To demonstrate that transformation helps in problem solving, several examples of problems solved using the congruence transformations are presented below. To remind the reader of the basic properties of the congruence transformations in a plane, as well as to explain the terminology used, we include a brief summary on each of the congruence transformations used in the several paragraphs that follow.

A congruence transformation T is a mapping in which to every point X of a plane just one point X' is determined, called the image of the point X in the transformation T, and in which, if X' and Y' are the images of two points X and Y in T, then the length of the segment $\overline{X'Y'}$ is equal to the length of the original segment $\overline{XY}$. In the case where a point X coincides with its image X' in T (i.e., $X \equiv X'$), we call the point X a "fixed (or invariant) point" in the transformation T. Similarly, if a line p is mapped by T onto itself (so that $p \equiv p'$), we say that p is a "fixed (or invariant) line" in T.

An image p' of a line p in the transformation T can always be constructed by joining the images of two distinct points of the line p—that is, $(X, Y) \in p$ implies $(X', Y') \in p'$.

The fact that the circle k has the center S and the radius r will be denoted by $k = (S, r)$. An image of such a circle in a congruence transformation T is a circle $k' = (S', r)$ of the same radius r, the center S' of which is the image of the center S of the circle k.

The inverse transformation T^{-1} to the transformation T maps every point X of the plane to the point Y, the image Y' of which in T coincides with X.

The most trivial case of the congruence mapping is the identity transformation, mapping every point of a plane to itself. All the points of the plane are therefore invariant under the identity transformation.

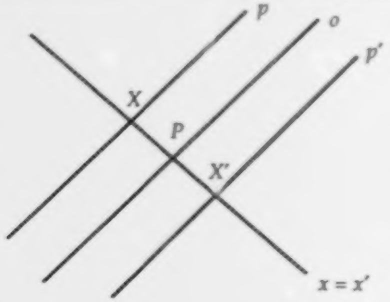


Figure 14.1

Axial Symmetry

Axial symmetry is a congruence transformation that has infinitely many fixed points, all lying on one line, called the "axis of symmetry." The axial symmetry is uniquely determined by its axis.

An image p' of a line p parallel to the axis o is also parallel to the axis o (i.e., $p \parallel o \parallel p'$).

All the lines perpendicular to the axis o are fixed lines, but the axis of symmetry is the only line every point of which is a fixed point.

An image of the point X lies on the line x, passing through the point X and perpendicular to the axis of symmetry. Denoting the intersection of the line x with the axis of symmetry by P, we can say that the image X' of the point X is exactly that point of the line x for which P is the midpoint of the segment $\overline{XX'}$ (see Figure 14.1).

The inverse transformation T^{-1} to an axial symmetry T is again an axial symmetry equal to T.

Rotation

Rotation is a congruence transformation with at least one fixed point, called the "center of rotation." The rotation is uniquely determined by its center and by a magnitude of an oriented angle, often called the "angle of rotation." The angle measurement is traditionally so defined that the magnitude of an angle is positive if it is counterclockwise oriented and is negative if an angle is clockwise oriented.

If S denotes the center of a rotation T and ω is the magnitude of the angle of rotation, an image X' of a point X is so defined that the magnitude of $\angle XSX'$ is equal to ω and $SX = SX'$ (see Figure 14.2).

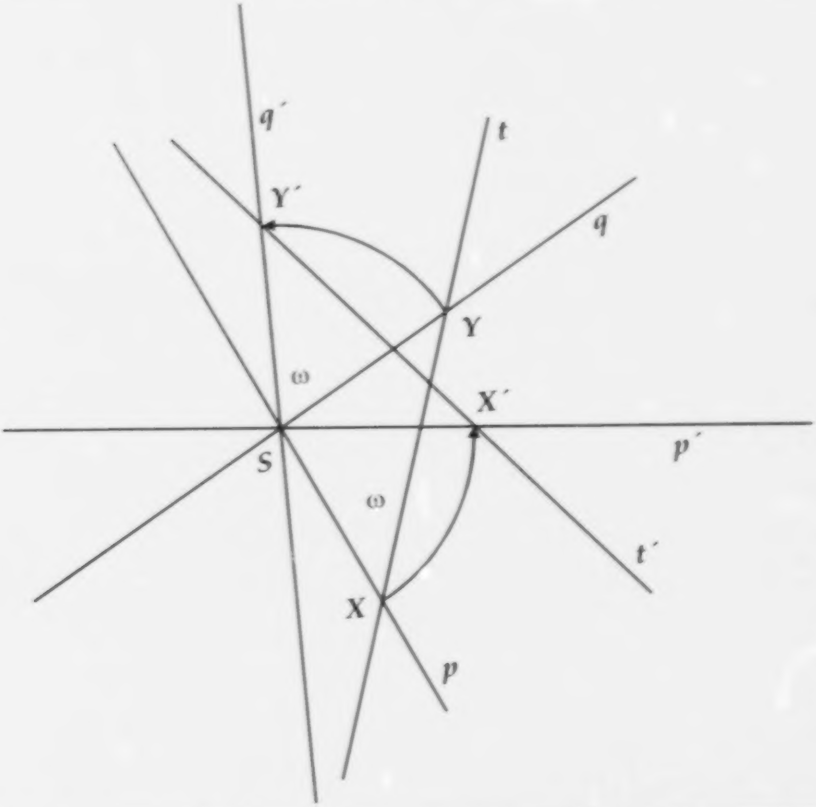


Figure 14.2

In the general case, there is just one fixed point—the center of rotation—and there are no fixed lines. In the case, though, where the absolute value of the angle of rotation is equal to π, the rotation has some special properties. The center of this rotation is called the "center of symmetry," and a rotation of this kind is called the "central symmetry." While the center of symmetry is still the only fixed point, all the lines passing through it are fixed lines. In addition, every line p is parallel to its image p' in the central symmetry (see Figure 14.3).

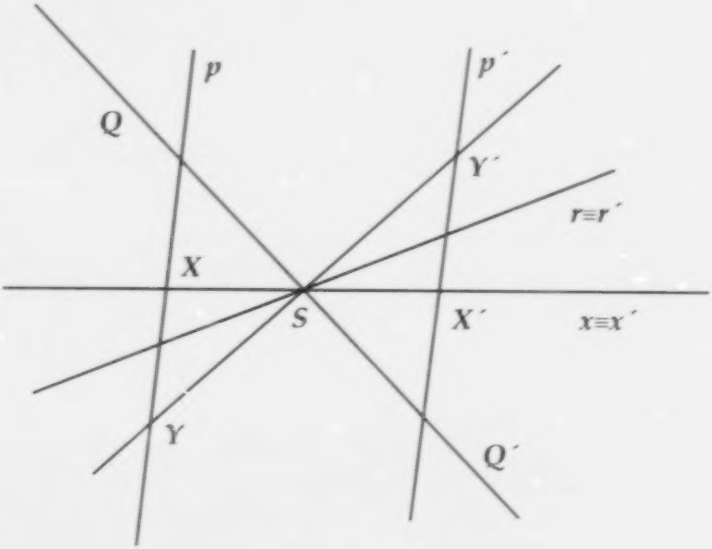


Figure 14.3

In the case in which the absolute value of the angle of rotation is equal to 2π, the rotation turns into the identity transformation, with all points, and therefore also with all lines, being invariant.

If the rotation T is given by its center S and by the angle of the magnitude ω, then the transformation T^{-1} inverse to T is again the rotation with the center S and the angle of rotation of the magnitude $-\omega$. Every central symmetry is inverse to itself: $T \equiv T^{-1}$.

Translation

Translation is a congruence relation determined by the oriented segment or, equivalently, by a vector. The length of the segment (or vector) is called the "magnitude of the translation," and the oriented direction of the segment (or vector) is called the "direction of the translation." If X' is an image of a point X in the translation given by the oriented segment $\overline{AB}$, the segments $\overline{XX'}$ and $\overline{AB}$ are of the same oriented direction and of the same length (see Figure 14.4).

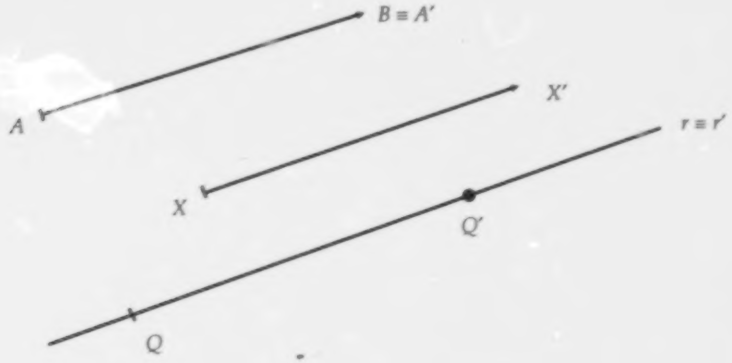


Figure 14.4

In the case where the segment defining the translation is of zero length, the translation is the identity transformation. Any translation that is not the identity transformation maps every line p to a line parallel to p, the lines of the direction of the translation are fixed lines, and there are no fixed points of this translation.

If the translation T is given by the oriented segment $\overline{AB}$, then the translation T^{-1} inverse to T is determined by the oriented segment $\overline{BA}$.

As an example of a problem from the abovementioned category, let us consider one whose solution demonstrates the use of a rotation.

Problem 1

GEOM

Two distinct lines p and q are given, and a point S. Draw a square $ABCD$ that satisfies the following conditions:

(a) Point S is the center of the square.
(b) The vertex A of the square lies on the line p.
(c) The vertex B, the counterclockwise neighbor-vertex of A, lies on line q.

Solution

① In the search for a solution, we assume first of all that the given problem really has some solution, and we sketch its picture in Figure 14.5. From now on, we shall be working backwards while looking for the relationships between the known and unknown (but desirable) objects (in most cases, unknown points).

⑥

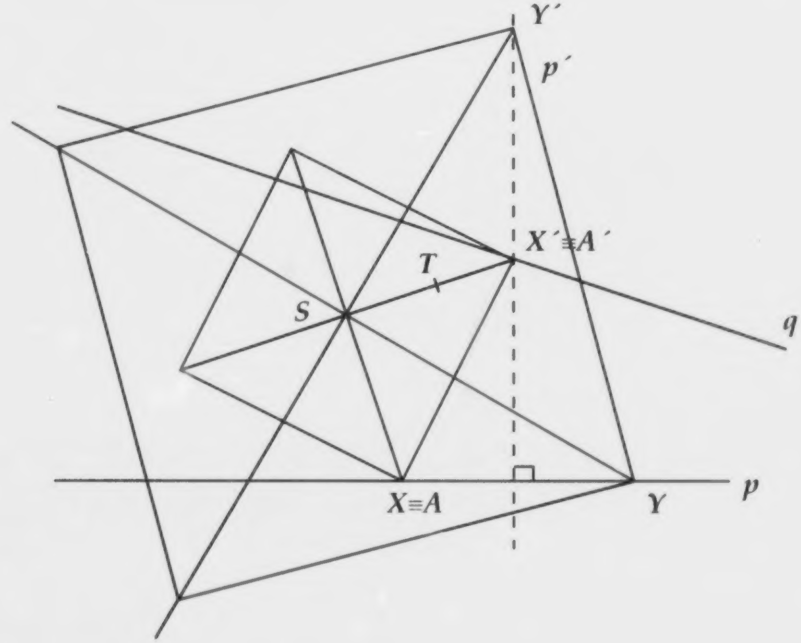


Figure 14.5

It can be observed that taking an arbitrary point X, we have already determined one square having one vertex in the chosen point X and the center in the given point S. The counterclockwise neighbor-vertex to X (let us denote it X') can be found on the ray $\overrightarrow{ST}$, for which $\angle XST$ is an oriented angle, equal to $+\pi/2$. There is always just one ray $\overrightarrow{ST}$ of this kind and just one point X' on it, such that $SX' = SX$.

The point X' is actually the image of X in the rotation R of the magnitude $+\pi/2$ about the center S. Taking another point Y on the line p, we can determine another square with the center in S and with the counterclockwise neighbor-vertex to Y in a point Y', which will appear again as an image of the chosen point Y in the mentioned rotation R. Joining both points X' and Y', we get line p'—an image of the line p in R—and since R is of a magnitude $+\pi/2$, p' is perpendicular to p. Thus, if A is a point on the line p, its image A' also lies on the line p', and $A' \equiv B$ is a counterclockwise neighbor-vertex to the vertex A of the square $ABCD$ for which we are looking.

Let us now consider the third condition of our problem, requiring vertex B to lie on the given line q. The above consideration implies that the point B ($B \equiv A'$), lying on the line p' and on the line q at the same time, is an intersection point of those two lines, $B = p' \cap q$.

We have already gathered enough information to be able to draw the desired square $ABCD$ when S, p, and q are given.

Construction

1. Draw the image of the line p in rotation R (in point S_1 with magnitude $\frac{\pi}{2}$). (See Figure 14.6):
   - (a) Construct the image X' of an arbitrary point $X \in p$ in the rotation R.
   - (b) Draw a line p' through X' perpendicular to p.
2. Denote the intersection point $B = p' \cap q$.
3. Find the point $A \in p$ as a point that is mapped by the rotation R into B (or as an image of the point B in the inverse rotation R^{-1} to R).
4. Complete the construction of the square $ABCD$, when its vertices A and B and the center S are known.

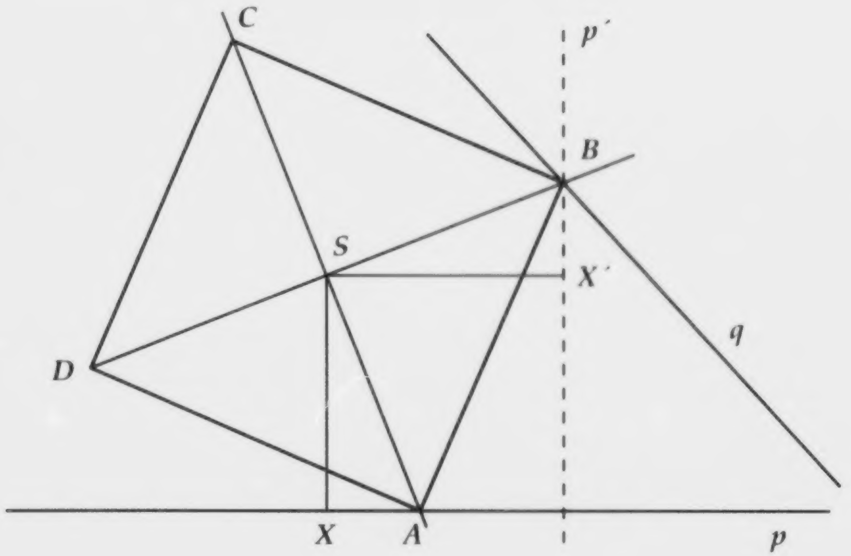


Figure 14.6

Accounting for All the Possibilities

⑪ When constructing the desired square, we can accomplish every step of the construction in a unique way, with the exception of step 2: The first step yields a unique line p' as an image of the given line p.

The second step, though, can result in one, none, or infinitely many points. We are getting a single point B in case the lines p' and q are not parallel, to which again just one point $A \in p$ exists, mapped by R to B and the construction results in the unique solution.

If the lines p' and q are parallel and distinct, they have no common point. No point B means no point A and, therefore, no solution of the problem.

If the lines p' and q are identical, then any point of the line $q \equiv p'$ can be taken for B, and every such point B provides us with one of infinitely many squares satisfying the conditions of the given problem.

The third condition of our problem, slightly altered, changes our problem into another one which provides the opportunity to employ a central symmetry (or, if the reader prefers, a rotation of the magnitude π) to get a solution easily.

Problem 2

GEOM

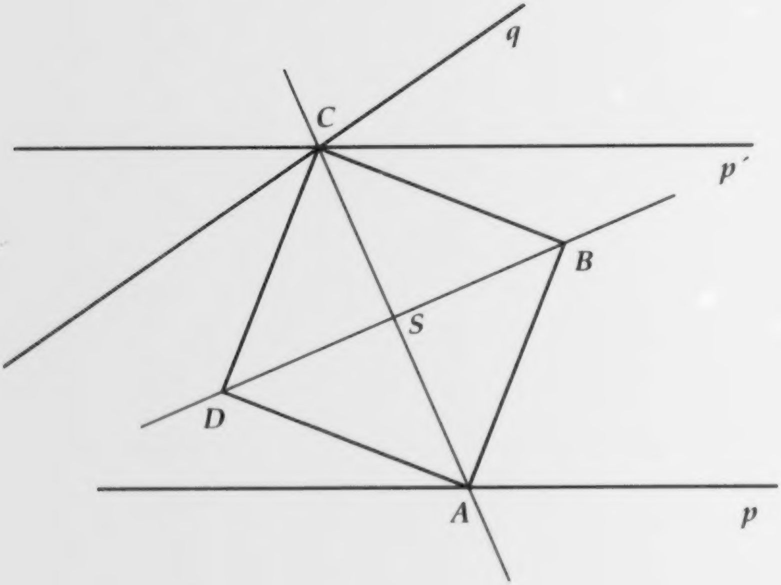
Two distinct lines p and q and a point S are given. Draw a square $ABCD$ that satisfies the following conditions:

(a) Point S is the center of the square.
(b) The vertex A of the square lies on the line p.
(c) The vertex C opposite to the vertex A lies on the line q.

Solution

① The solution to this problem is very similar to that of problem 1. Provided that a square $ABCD$ complying with the three conditions exists (such as the one sketched in Figure 14.7), the vertices A and C are symmetric with respect to the center S.

Because the vertex A lies on the line p (to satisfy the second condition of the problem), the vertex C lies on the image p' of the line p in the same symmetry (let us denote it S) in which A and C form a pair of the corresponding points. The key point here is the vertex C; we know that it must lie on the line q (according to the third condition) as well as on the line p', and therefore $C = q \cap p'$. Having the position of the vertex C established, we can find its image A in the symmetry S as



⑥

Figure 14.7

an intersection point $A = \overleftrightarrow{CS} \cap p$ and complete the construction of the square *ABCD* without further difficulty.

Construction

We can consider the given lines p and q, and the given point S, drawn already in our plane. (See Figure 14.8)

1. Draw the line p' as an image of p in the symmetry S (reflection in point S).
   (a) Draw the image X' of an arbitrary point $X \in p$ in the symmetry S.
   (b) Through the point X' draw the line $p' \parallel p$.
2. Denote $C = p' \cap q$.
3. Draw the image A of the point C in the symmetry S.
4. Complete the construction of the square *ABCD* when the vertices A and C and the center S are given.

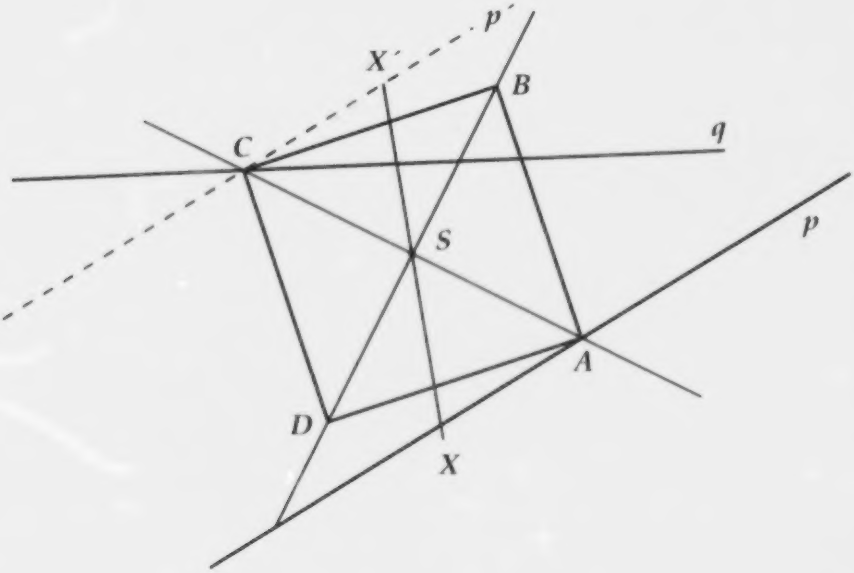


Figure 14.8

Accounting for All the Possibilities

⑪ In an analogy with the previous problem, we can conclude that once the point C exists (the lines q and p' are not parallel), a unique resulting square $ABCD$ is the only solution, while there is no solution when the lines q and p' are parallel and distinct, and there are infinitely many solutions in the case of the coinciding lines $q \equiv p'$. In the last case, any point of the line $q \equiv p'$ can be taken for a point C and used as the vertex of square $ABCD$ with the center in S and the vertex A lying on the given line p.

The character of the next problem suggests the use of a reflection in a line because the geometrical object we are looking for is required to be symmetrical with respect to the given line.

Problem 3

Given two lines p and q and a circle $k = (S, r)$ with the center S and the radius r. Draw an isosceles ΔABC having $C \in p$, $A \in k$, and $B \in q$, for which $CA = CB = d$ is a given distance and $p \perp \overline{AB}$.

GEOM

Solution

Draw a picture and determine characteristics of objects while working backwards. ①

Let us suppose that the problem has some solutions, one of which looks like the one sketched in Figure 14.9.

⑥

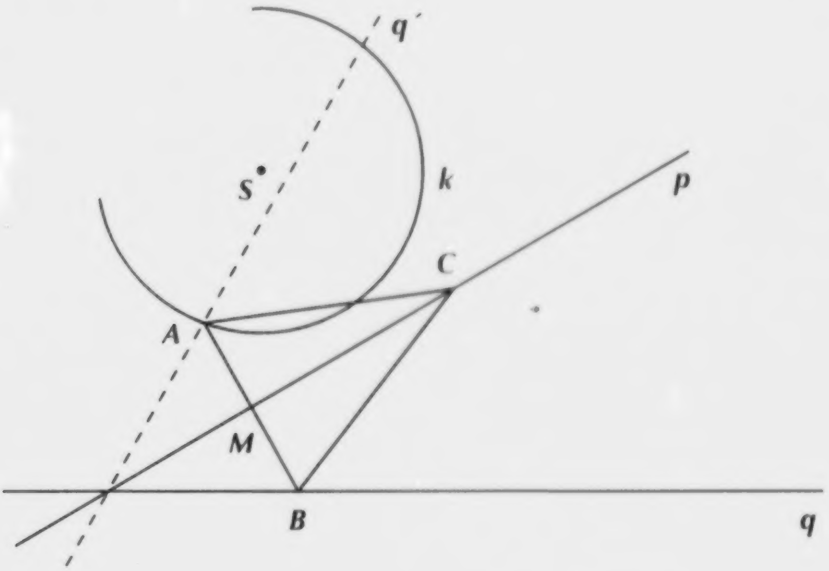


Figure 14.9

Since $\overline{AB}$ is perpendicular to p and $CB = CA$, the line p intersects $\overleftrightarrow{AB}$ in the midpoint M of $\overline{AB}$. $\overline{CM}$ is then the altitude, the median, the axis of the side $\overline{AB}$, and the axis of $\angle C$ of ΔABC at the same time. The points A and B are therefore mapped on each other by the reflection R in the line p.

Because the point B lies on the line q, the point A must lie where the images of all the points of the line q lie—that is, on the image q' of the line q in the reflection R. Since A is required to lie also on the given circle k, it is a point of intersection of k with q'. The point B can be found on the line q as that point of q that is mapped by R on A. The last vertex C of ΔABC, having the distances to A and B equal ($CA = CB = d$), lies on a circle $m = (A, d)$ with the center A and the radius d (and also on a circle with the center B and the same radius d).

Construction

Lines p and q and a circle k are given in the plane; d is a given distance. (See Figure 14.9)

1. Draw the image q' of the line q in the reflection R through line p.
2. Denote $A = k \cap q'$.
3. Draw a circle $m = (A, d)$.
4. Denote $C = m \cap p$.
5. Find B as an image of A in the reflection R.
6. Draw the triangle ΔABC.

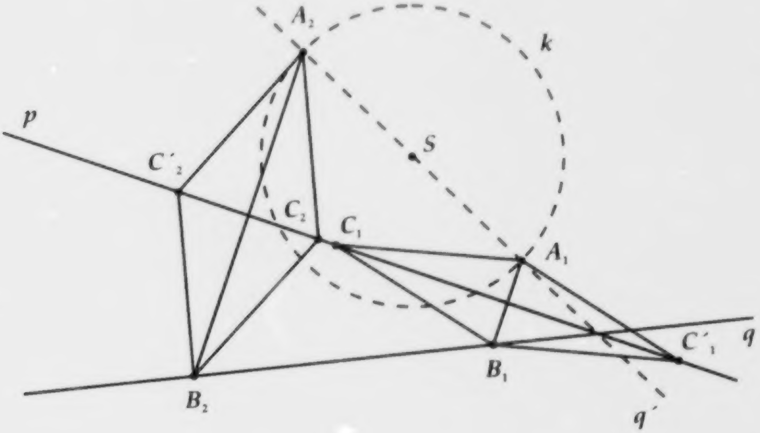


Figure 14.10

Accounting for All the Possibilities

Let us look through the construction again, surveying each step in order to find out whether that particular step can always be accomplished and, if this is the case, in how many ways it can be done. ⑪

The first step can certainly always be realized, and the resulting line q' is unique.

In the second step, looking for the points common to line q' and the circle k, we can find two distinct intersection points (A_1, A_2), a single point in which the line q' touches the circle k, or no common point at all.

In the third step, therefore, we can have two circles $m_1 = (A_1, d)$ and $m_2 = (A_2, d)$, only one of them, or none.

In the fourth step, every circle m (if there are any) can have, in general, two distinct points, one point, or no point in common with the line p. Thus, the fourth step can result in up to four distinct points in the most favorable case, shown in Figure 14.10, where they are denoted as C_1, C_2, C'_1, C'_2.

To every point A of the base of ΔABC there exists just one point B constructed in the fifth step of our construction as an image of A in the reflection R.

In step six, we can always construct ΔABC except in the case in which the points A, B, and C are collinear, which will be the case when C is the point of tangency of the line p and the circle m.

Summarizing the results, we can say the following:

a. There are four distinct solutions in the case in which step 2 results in two distinct points and step 4 yields four distinct points (see Figure 14.10).
b. There are two solutions in the case in which
    (i) only one point A was found in step 2 and the single circle $m = (A, d)$ intersects the line p in two distinct points (C, C') or
    (ii) there are two points A_1, A_2 provided by step 2; but of the two circles $m_1 = (A_1, d)$ and $m_2 = (A_2, d)$, only one intersects the line p in two distinct points.
c. In all other cases, there is no solution.

Problem 4

Draw a triangle ΔABC, if its side $c = AB$, its median t_a, and an acute angle $\omega = \angle(t_a, b)$ between its median t_a and its side b are given. GEOM

Solution

① Let us suppose that such a triangle really exists and that it looks like the one sketched in Figure 14.11. Let us denote by S_A the midpoint of the side a. Then $t_a = AS_A$.

⑥

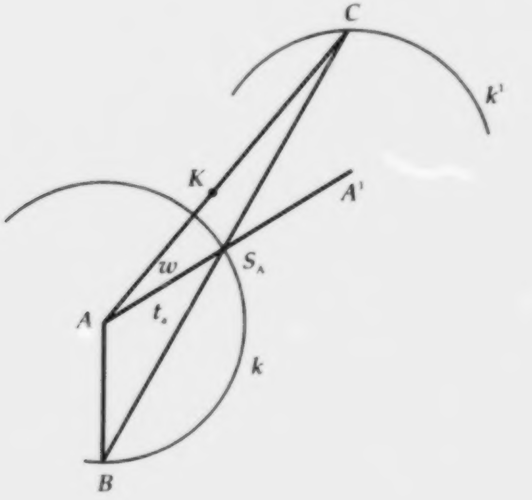


Figure 14.11

Points B and C are symmetrical with respect to the point S_A, and since B lies on the circle with the center A and the radius c (which we shall refer to as k), we can look for the point C on the circle k', which is the image of the circle k in the already mentioned central symmetry S with the center S_A. The point C is also the endpoint of the side b, the angle of which with the median t_a is known to be ω.

Thus, if the angle $\omega = \angle(b, t_a)$ is positioned as $\angle KAS_A$, the point A' found as an image of the point A in the symmetry S, and the circle $k' = (A', c)$ drawn, the point C is an intersection point (if it exists) of the circle k' and the side $\overrightarrow{AK}$ of the angle ω.

Construction

1. Construct $\angle KAS_A$ of the magnitude ω with a point K on one side of the angle and a point S_A on the other side, so that the length of $\overline{AS_A}$ is equal to t_a. (See Figure 14.12)
2. Draw the circle $k = (A, c)$.
3. Construct an image A' of the point A in the central symmetry S with the center S_A.
4. Draw an image $k' = (A', c)$ of the circle k in S.
5. Denote $\{C_1, C_2\} = \overrightarrow{AK} \cap k'$ if they exist.
6. Denote by B_1 and B_2 the points of the circle k that are mapped by the symmetry S to the points C_1 and C_2, respectively. Notice that B_1 and B_2 are also images of C_1 and C_2 in the symmetry S, because $S \equiv S^{-1}$.
7. Draw the resulting triangles ΔAB_1C_1 and ΔAB_2C_2.

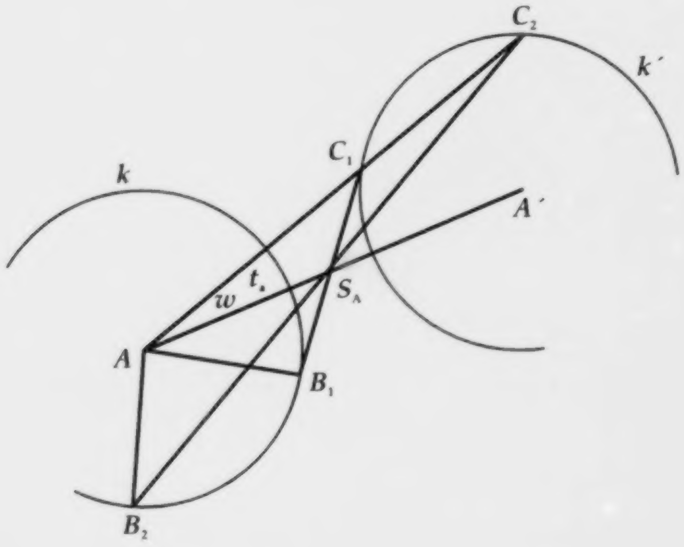


Figure 14.12

⑪ From the above discussion it follows that there are two solutions of the problem when $\overrightarrow{AK}$ intersects the circle k' in two distinct points, as in Figure 14.12. There is just one solution when $\overrightarrow{AK}$ is tangent to the circle k', while there is no triangle with the given properties if $\overrightarrow{AK}$ has no common points with the circle k'.

Problem 5

GEOM

Two circles $k_1 = (S_1, r_1)$ and $k_2 = (S_2, r_2)$ and a pair of the distinct points A and B are given. Draw a segment $\overline{XY}$ parallel to $\overline{AB}$, the endpoint X of which lies on k_1, the other endpoint Y of which lies on k_2 and $XY = AB$.

Solution

① Assuming again that a solution exists for the given circles k_1 and k_2 and the pair of points A and B, let us sketch one of the possible configurations (Figure 14.13).

⑥

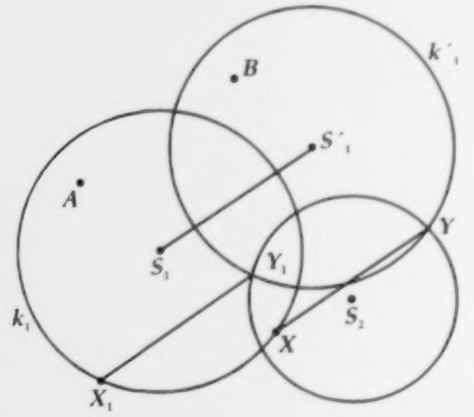


Figure 14.13

The point Y can be considered as being the image of the point X in the translation T determined by the oriented segment $\overline{AB}$. While the point X runs on along the circle k_1, the point Y traces another circle $k'_1 = (S'_1, r_1)$, the image of k_1 in the translation T. The circles k_1 and k'_1 are congruent, and the center S'_1 of the circle k'_1 is the image of S_1 in T. The point Y can be found as an intersection point of the two circles k_2 and k'_1, and the point X is the point of the circle k_1 mapped by T on Y.

There can be up to two intersection points of k_2 and k'_1 (as in Figure 14.13), yielding two distinct solutions. Since $\overline{YX}$ (with the opposite orientation) where $X \in k_1$ and $Y \in k_2$ is also a solution of the given problem, the point Y can also be taken as an image of X in the inverse translation T^{-1}, lying on a circle $k''_1 = (S''_1, r_1)$, which is an image of k_1 in the translation T^{-1}. Two possible distinct intersection points of the circles k_2 and k''_1 can provide for two more solutions of the problem.

Construction

Let us consider that the circles k_1 and k_2 and the points A and B are already positioned in the plane.

1. Draw the images S'_1 and S''_1 of the center S_1 in the translations T and T^{-1} respectively, where T is determined by the oriented segment $\overline{AB}$ and T^{-1} by $\overline{BA}$ (of the opposite orientation).
2. Draw the circles $k'_1 = (S'_1, r_1)$ and $k''_1 = (S''_1, r_1)$.
3. Determine the intersection points $\{Y_1, Y_2\} = k_2 \cap k'_1$, $\{Y_3, Y_4\} = k_2 \cap k''_1$, if they exist.
4. Find the points X_1 and X_2 as the points mapped by T to Y_1 and Y_2, respectively, and the points X_3 and X_4 as the points mapped by T^{-1} to Y_3 and Y_4, respectively.
5. Draw the resulting segments and $\overline{X_1 Y_1}$, $\overline{X_2 Y_2}$, $\overline{X_3 Y_3}$, and $\overline{X_4 Y_4}$.

Accounting for All the Possibilities

In the most favorable case (shown in Figure 14.14), we can have four distinct ⑪
resulting segments, when each of the pairs of the circles k_2, k'_1 and k_2, k''_1 has two distinct intersection points. In the less favorable cases, there can be three, two, one, or no intersections, resulting in three, two, one, or no solution segments.

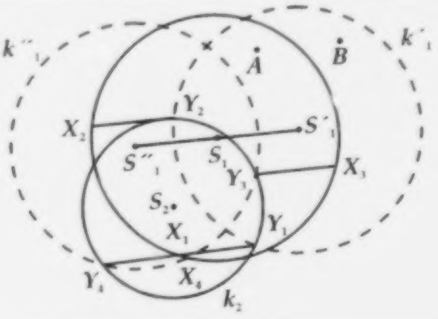


Figure 14.14

Problem 6

GEOM

Given two distinct lines p and q, a circle $k = (S, r)$, a point C, and a distance d. Draw a triangle ΔABC, such that $A \in k$, $B \in q$, $\overline{AB} \parallel p$, and $AB = d$.

Solution

① We assume again that our problem has at least one solution, the picture of which is shown in Figure 14.15. Let us investigate relationships between the given objects and those we need in order to be able to construct the resulting triangle.

Because we know the length of the side AB ($= d$) of ΔABC, as well as the direction of the segment $\overline{AB}$ ($\overline{AB} \parallel p$), we can take the vertex B for an image of the vertex A in a translation T of the magnitude d and of the direction given by $\overrightarrow{AB}$. But also, on the contrary, the vertex A can be taken for an image of the vertex B in the translation T^{-1} (inverse to T) of the same magnitude d and of the opposite direction, given by the opposite ray $\overrightarrow{BA}$.

Let us consider, in the first case, the use of the translation T. Provided the point A lies on the circle k, point B has to lie where all the images in T of all the points of the circle k lie—that is, on the image of the circle k in the translation T. This, as we already know, is again a circle $k' = (S', r)$, where S' is the image of the point S in T. Since our problem requires the point $B \equiv A'$ to lie on the given line q, we can find B as the intersection point of this line with the circle k'. Two such points of intersection can, in general, be found, and still another two may be constructed as intersection points of the image $k'' = (S'', r)$ of the circle k in the inverse translation T^{-1} with the given line q.

⑥

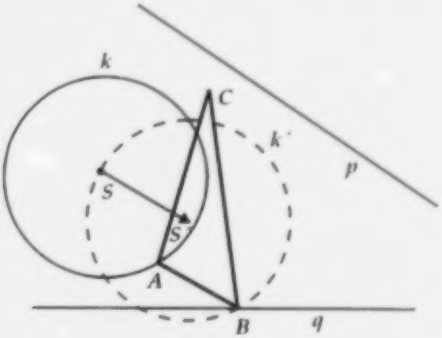


Figure 14.15

Once the vertex B is determined, we shall find the last vertex A of ΔABC as the point of the circle k mapped by the corresponding translation to the point B.

Construction

Let us suppose that the given objects p, q, $k = (S, r)$, and C are positioned in a plane and a non-zero distance d is known. Then the desired triangle can be constructed in the following way:

1. Draw an image $k' = (S', r)$ of the circle k in the translation T, where T has magnitude d on direction $\overrightarrow{AB}$. (See Figure 14.15)
2. Draw an image $k'' = (S'', r)$ of the circle k in the translation T^{-1}.
3. Denote $\{B_1, B_2\} = k' \cap q$ and $\{B_3, B_4\} = k'' \cap q$.
4. Determine the points A_1, A_2, A_3, A_4 of the circle k mapped to the points B_1, B_2, B_3, B_4, respectively.
5. Draw the triangles $\Delta A_1B_1C_1$, $\Delta A_2B_2C_2$, $\Delta A_3B_3C_3$, $\Delta A_4B_4C_4$, if they exist.

Accounting for All the Possibilities

Already in the discussion of our problem we have accounted for all the solutions possible, and we have found that in the most favorable configuration we ⑪
can have four solutions, when the given line q intersects both of the circles k' and k'' in two distinct points, as is the case in Figure 14.16.

All the other configurations of the given elements can result in three, two, one, or no solutions, depending on the relative positions of the line q and the circles k' and k''.

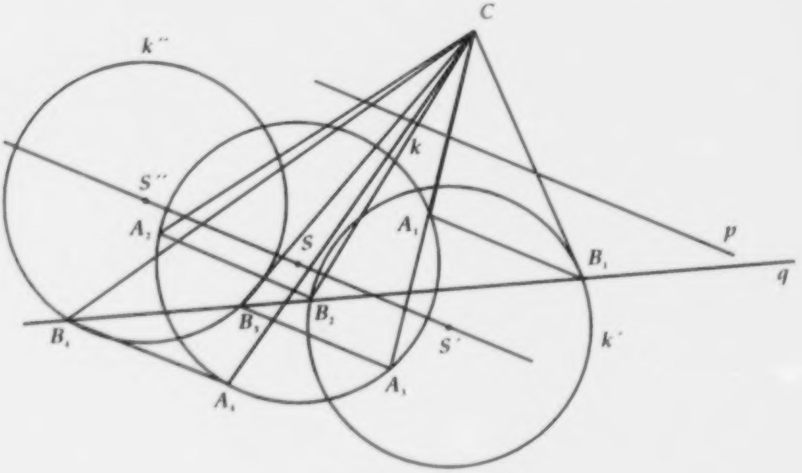


Figure 14.16

Problem 7

GEOM

When a line p, two distinct circles $k_1 = (S_1, r_1)$ and $k_2 = (S_2, r_2)$, and a distance d are given, draw the line parallel to p cutting the circles k_1 and k_2 in the chords $\overline{X_1Y_1}$ and $\overline{X_2Y_2}$, respectively, in such a way that the sum $X_1Y_1 + X_2Y_2$ is of the given length d.

Solution

① Provided that a solution of the problem exists, we can assume that the configuration looks like the one in Figure 14.17.

Let us consider a case when the circle k_2 is so shifted in the direction of the line p that the two points Y_1 and X_2 coincide. Let us denote this shifted position of the circle k_2 by k_2^+ and its center by S_2^+. Then the shifted position of the point $Y_2 \in k_2^+$ can be taken for the image X'_1 of the point $X_1 \in k_1$ in a translation T determined by the direction of the line p (and of whatever orientation) and by the length d. The point X'_1 has to lie on an image $k'_1 = (S'_1, r_1)$ of the circle k_1 in the translation T. The center S_2^+ of the circle k_2^+ lies on the axis o of symmetry of the shifted chord $\overline{X_2Y_2}$, which is the axis of symmetry of the segment $S_1S'_1$ at the same time. The circle $k_2^+ = (S_2^+, r_2)$ can therefore be drawn as an image of k_2 in the translation T^+ determined by the distance of the center S_2 from the axis o in the direction parallel with p (i.e., perpendicular to o) and orientation from S_2 to the axis o.

Thus, the point Y_1 (if it exists) can be found as an intersection $k_1 \cap k_2^+$. The line through Y_1 parallel to p may intersect k_1 in another point X_1, the circle k_2^+ in another point X'_1 (which is also the point of k') and the circle k_2 in the points X_2 and Y_2. Since the chord $Y_1X'_1$ of k_2^+ is the image of the chord $\overline{X_2Y_2}$ of the circle k_2, they are of the same length and so $d = X_1Y_1 + Y_1X'_1 = X_1Y_1 + X_2Y_2$.

⑥

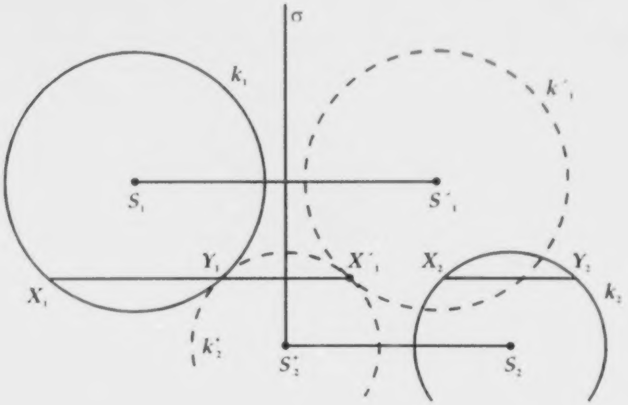


Figure 14.17

Construction

The given elements p, k_1, and k_2 are placed in the plane, and the distance d is known.

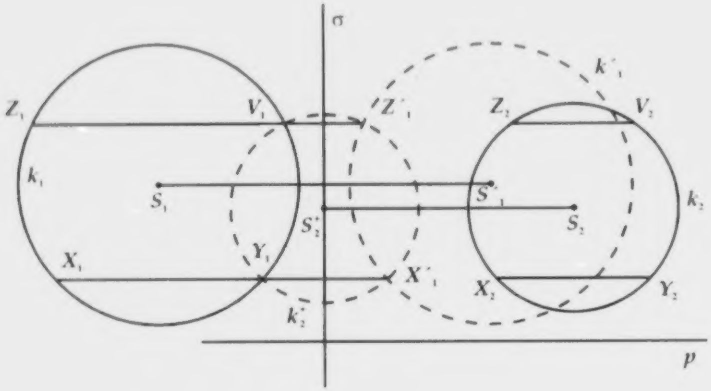


Figure 14.18

1. Draw the image S'_1 of the center S_1 in the translation T and the image circle $k'_1 = (S'_1, r_1)$ of the circle k_1.
2. Draw the axis o of the segment $\overline{S_1S'_1}$.
3. Construct the point $S_2^{\ddagger}$ such that
   (a) $S_2^+ \in o$
   (b) $\overline{S_2^+ S_2} \parallel p$.
4. Draw the circle $k_2^{\ddagger} = (S_2^{\ddagger}, r_2)$.
5. Denote $Y_1 = k_1 \cap k_2^{\ddagger}$.
6. Draw the line q through Y_1 parallel to p.
7. Denote the points X_1, Y_1, X_2, and Y_2 for which $\{X_1, Y_1\} = q \cap k_1$ and $\{X_2, Y_2\} = q \cap k_2$.

Accounting for All the Possibilities

The first four steps of the construction can always be accomplished just one way. ⑪

The fifth step, though, can result in two, one, or no points Y_1. In Figures 14.18 through 14.21, we have two distinct intersections $\{Y_1, V_1\} = k_1 \cap k_2^+$. The line q through Y_1 and the line q' through V_1 yield two pairs of chords, $\overline{X_1Y_1}$, $\overline{X_2Y_2}$ and $\overline{Z_1V_1}$, $\overline{Z_2V_2}$, respectively.

In step 7, some intersections of q and/or q' may coincide. Thus, in Figure 14.19, one chord of the circle k_2 is of zero length, while in Figure 14.20, the chord

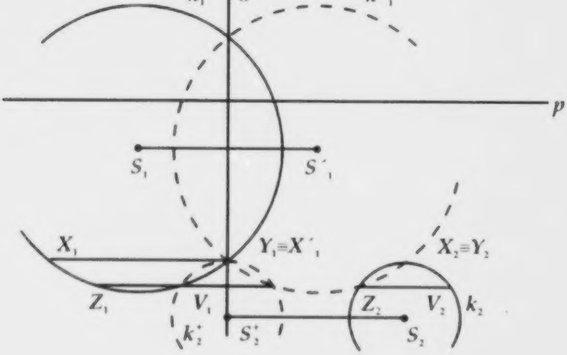

k_1
σ
k'_1
p
S_1
S'_1
X_1
$Y_1 \equiv X'_1$
$X_2 \equiv Y_2$
Z_1
V_1
k^*_2
S^*_2
Z_2
V_2
k_2
S_2

Figure 14.19

of zero length is one of the circle k_1, and in Figure 14.21, both of the chords of the circle k_2 have lengths equal to zero.

Step 5 results in one solution in Figure 14.17 and in no solution in Figure 14.22. It can be observed that (besides some other, more intricate cases) there is always no solution in the case in which $r_1 + r_2 < d/2$ or the line p is of such a direction that no line q parallel to p can be found intersecting both of the circles k_1 and k_2 at the same time.

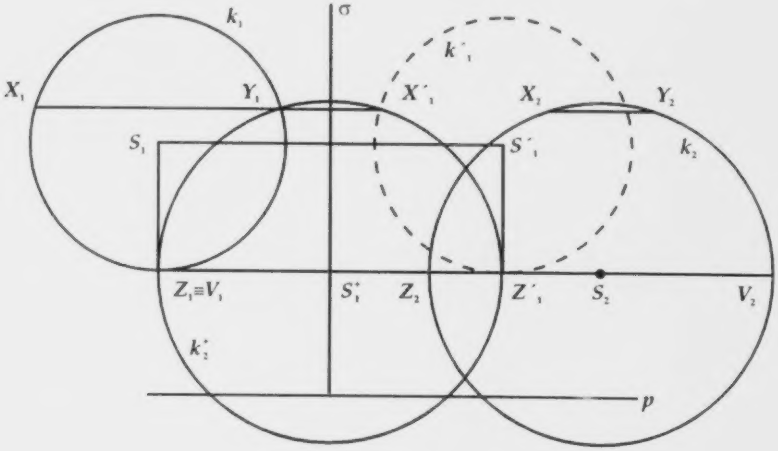

k_1
σ
k'_1
X_1
Y_1
X'_1
X_2
Y_2
S_1
S'_1
k_2
$Z_1 \equiv V_1$
S^*_1
Z_2
Z'_1
S_2
V_2
k^*_2
p

Figure 14.20

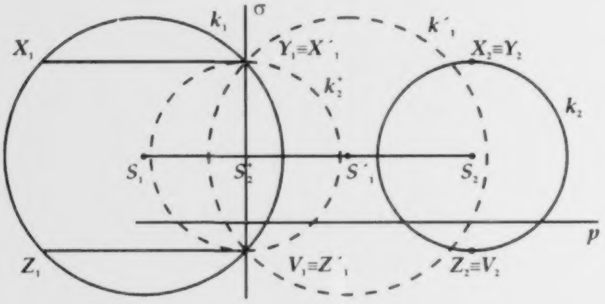

σ
k_1
k'_1
X_1
$Y_1 \equiv X'_1$
$X_2 \equiv Y_2$
k_2^*
k_2
S_1
S_2^*
S'_1
S_2
p
Z_1
$V_1 \equiv Z'_1$
$Z_2 \equiv V_2$

Figure 14.21

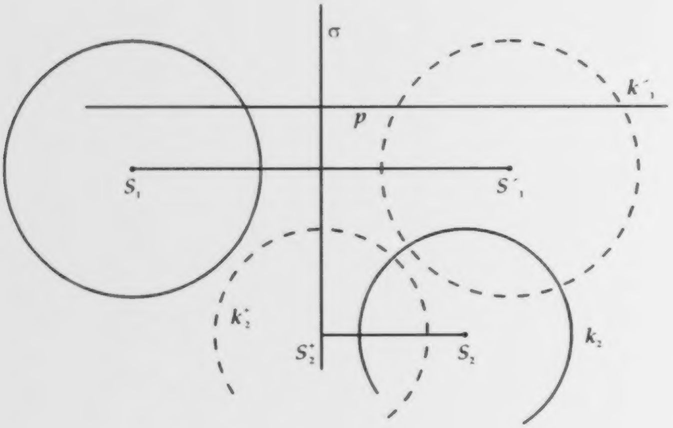

σ
k'_1
p
S_1
S'_1
k_2^*
S_2^*
S_2
k_2

Figure 14.22

Final Remarks on Congruence Transformations

The general subject of geometry was characterized for the first time by German mathematician Felix Klein in his *Erlanger's Program,* published in 1872. The subject of Euclidean geometry, in particular, then appeared to be the investigation of those properties of figures in a Euclidean space that do not change when a figure is transformed by any congruence transformation (also called "isometry") of the space. Since a congruence transformation preserves the shape as well as the size of a transformed figure, the properties—such as all the incidence relations (e.g., a point-line incidence or intersection of lines), order relations (of points on a line or lines in a pencil), perpendicularity, parallelism, and the magnitudes of segments and angles (and consequently also lengths of curves, areas of surfaces, and volumes of bodies)—are studied in Euclidean geometry.

To understand the key role of the congruence transformations in geometrical problem solving, let us have a look now at the question of the number of distinct solutions of a geometrical problem. We have to discern two categories of problems: The first is represented by our problem 4, while all of the remaining problems in this chapter belong to the second category. The distinction between the categories lies in the fact that there are no claims on the position of the object (triangle) in problem 4 but on its size and shape, while there are many claims in all the other problems on the position of the resulting objects (triangles or segments) in relation to the given points, lines, and/or circles already placed in a plane. From this point of view, we can call the problems of this category the "positional" problems, and the problems represented by problem 4, "nonpositional" problems.

There is no difficulty in determining the number of distinct solutions of the positional problem. The traditional wording "Find an object, the position of which satisfies given conditions" is generally understood to mean "Find all such objects, some of which may be, while the others need not be, congruent."

Not so in the case of nonpositional problems. To see the difference more clearly, imagine solving the problem consisting of construction of ΔABC, of which the lengths of sides a, b, c are given (Figure 14.23).

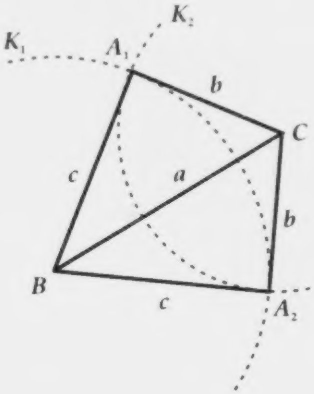


Figure 14.23

The first thing we do when constructing the triangle is choose the position of one side, say BC, by placing the points B and C somewhere in a plane so that

$BC = a$. In so doing, we are changing the original nonpositional problem into the positional one, the solution of which can be found by determining the last vertex A of the triangle by method of loci: the circle $k_1 = (B, c)$, being the locus of points the distance of which from the point B is equal to c, and $k_2 = (C, b)$, being the locus of points b units from the point C, intersect each other in two points A_1 and A_2, provided that ABC exists (i.e., that the triangle inequality $a < b + c$ holds for the lengths of its sides). The positional problem under consideration can only have two distinct solutions ΔA_1BC and ΔA_2BC, symmetrical in the line BC on which their common side $BC = a$ lies, since otherwise either A, B, and C are collinear points or there are no such points for which $AB = c$, $AC = b$, and $BC = a$.

From the point of view of the original nonpositional problem, it is obvious that if ΔABC is a solution of the problem, then any triangle ΔPQR to which ΔABC is mapped by an arbitrary congruence transformation (i.e., any triangle congruent with ΔABC) is another solution of this nonpositional problem; the lengths of all sides of both triangles are equal, and there are no claims on the position of the resulting triangles.

It can also be said that if ΔABC is one solution of a nonpositional problem, then the whole family of all the triangles congruent with ΔABC is a solution represented by the triangle ΔABC. On the other hand, two triangles are two distinct solutions of a nonpositional problem only if they are not congruent (as in problem 4), and therefore, they represent two disjoint families of congruent triangles.

The same principle in making the decision about whether or not two objects are two distinct solutions of the given nonpositional problem can also be interpreted in a more manipulative way, by handling the resulting objects cut out of cardboard. Two such objects can be considered distinct solutions of the problem only if they cannot be moved (without deformation) to coincide with one another.

A comprehension of the properties of congruence transformations in a plane is the necessary prerequisite to an understanding of the spatial congruence transformations, which, besides representing the most frequent motions in our physical world, are used in many branches of pure and applied mathematics, robotics, physics, mineralogy, and many other sciences.

Many other planar geometrical transformations are useful tools in problem solving. One of the most often used is homothety. Some of them (such as circular inversion and reciprocation) can be used to turn a difficult problem into an easier one. The given configuration is mapped by a suitable transformation to another one, where the solution is found. The image of this solution in the inverse transformation then solves the given problem. The detailed treatment of this method, however, is beyond the scope of the chapter. The interested reader can find excellent information on these methods in *Fundamentals of Modern Elementary Geometry*, by Howard Eves (Boston: Jones & Bartlett, 1992).

15

Graph Theory

Tools to Solve Mathematical Problems

WEI LEE

Introduction

The Swiss mathematician Leonhard Euler, who wrote a paper in 1736 solving the popular Königsberg Bridge problem, was the first mathematician to work in the area of graph theory. In recent years, graph theory has become an important area of mathematics where a great deal of research is being conducted. The current interest in graph theory among mathematicians is perhaps due to the fact that it has a large number of applications in both academic and business areas.

Although it can be a powerful tool, graph theory is not often suggested as a method for problem solving. In this chapter, I will try to offer some insight into using graph theory in problem solving, with the hope that it will be used more often.

AUTHOR'S NOTE: I would like to give my special thanks to Dr. Stefan A. Burr and Mrs. Bessie Burr, who put in a tremendous amount of time proofreading this chapter. Their corrections, comments, and encouragement made it possible for me to finish this chapter.

Graph theory has also played a very important role in mathematics contests. Some contest problems are graph theory problems in disguise. When you look at these problems, you may not know where to start. But if you familiarize yourself with some basic graph theory definitions and theorems and their applications, these problems will become more solvable.

This chapter consists of six sections. Section 1 introduces graph theory definitions. Section 2 presents selected problems that illustrate several strategies incorporating graph theory and problem solving. Sections 3 and 4 use these graph theory tools to solve the problems. Section 5 introduces directed graphs to solve problems involving contests. Section 6 provides additional exercises for readers to practice what they have just learned about graph theory.

1. Basic Terminology

The graphs in graph theory consist of points and lines that connect these points. The graph in Figure 15.1a consists of 8 points $v_1, v_2, \ldots, v_8$, and 9 lines $e_1, e_2, \ldots, e_9$. The graph in Figure 15.1b consists of 5 points $v_1, v_2, \ldots, v_5$, and 8 lines $e_1, e_2, \ldots, e_8$.

There is a difference between these two graphs. The one in Figure 15.1b has arrows on each line, representing the direction of the line. We call this kind of graph a *directed graph.* We call the graph in Figure 15.1a an *undirected graph.*

We use capital letters to represent graphs. Sometimes we add subscripts or superscripts—such as G_1, G_2, G^1, and so on—to represent graphs. In a directed graph, we assume that if there are two lines between two points, these two lines must have opposite directions. In Figure 15.1b, there are two lines with opposite directions between the points v_1 and v_2.

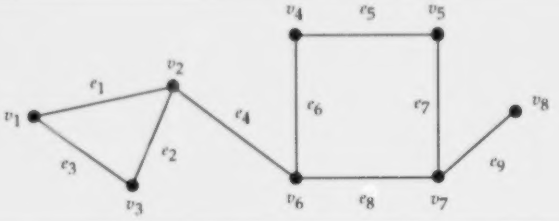


Figure 15.1a

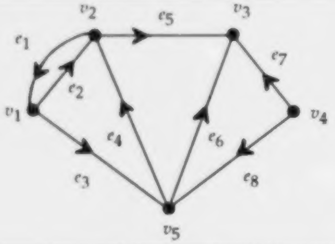


Figure 15.1b

The points in a graph are called *vertices*. We call lines in an undirected graph *edges*, and the lines in a directed graph *arcs*. The edges in a graph may be represented by writing both endpoints of that edge. For example, $v_1 v_2$ means the same as e_1 in Figure 15.1a.

The most important thing in a graph is its number of vertices. Questions such as whether the lines are straight or curved are not important. Let's look at graphs G_1 (Figure 15.2a) and G_2 (Figure 15.2b). It looks as if they are two different graphs, but actually the two graphs are isomorphic (i.e., they are the same graph). Both have six vertices ($v_1, v_2, v_3, v_4, v_5, v_6$), and exactly the same pairs of vertices are joined by edges.

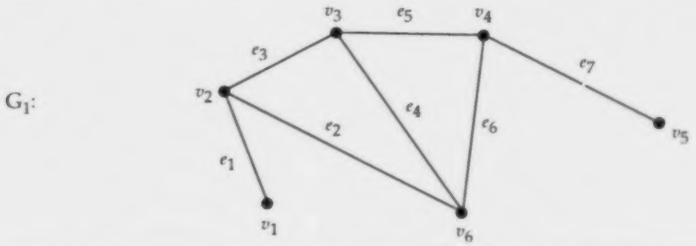


Figure 15.2a

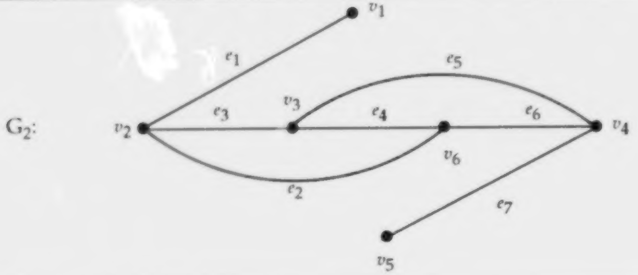


Figure 15.2b

The definition of some common terms used in graph theory will be helpful here. In an undirected graph, if an edge e_i connects the vertices v_j and v_k, then we say v_j and v_k are the *endpoints* of the edge e_i, and the vertices v_j and v_k are *incident* with the edge e_i. Moreover, we say v_j and v_k are *adjacent vertices*. Similarly, if two edges e_i and e_j share a common vertex, then we call e_i and e_j *adjacent edges*.

In an undirected graph, if there are edges between every two vertices in that graph, then we call it a *complete graph*. If the graph has n vertices, we denote the complete graph as K_n. Figure 15.3 shows the graphs K_2, K_3, K_4 and K_5. Note that in K_4 and K_5 the intersection of the two crossing lines is not a vertex in the graph.

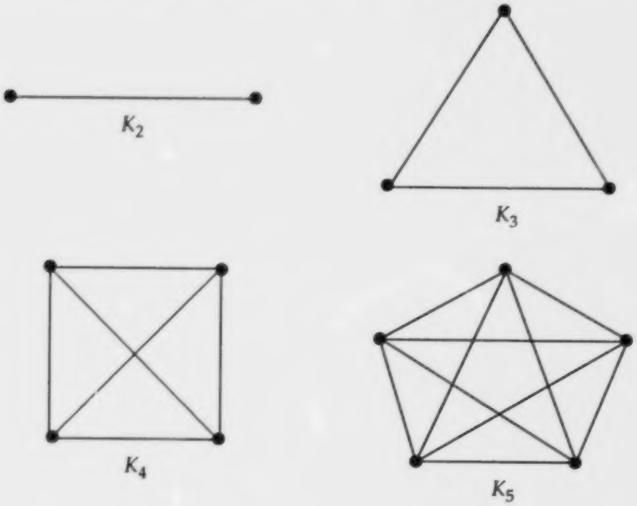


Figure 15.3

If v is a vertex of an undirected graph, we say the number of edges with v as their endpoint is the *degree* of the vertex v, denoted by deg (v). For example, in Figure 15.4, deg $(v_1) = 2$, deg $(v_2) = 3$, deg $(v_3) = 3$, deg $(v_4) = 1$, deg $(v_5) = 4$, deg $(v_6) = 1$, and deg $(v_7) = 0$.

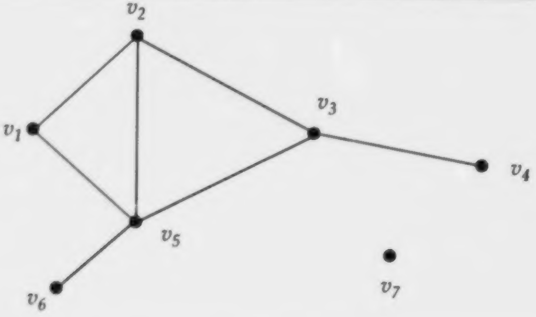


Figure 15.4

Vertices with degree 0, like v_7, are called *isolated vertices*. The following theorem is concerned with the degrees of vertices.

Theorem 1. In any graph, the sum of the degrees of the vertices equals twice the number of the edges.

Before we prove this theorem, let's see if the theorem applies to the graph in Figure 15.4. The sum of the degrees of the vertices of this graph is 2 + 3 + 3 + 1 + 4 + 1 + 0 = 14, and the number of edges in the graph is 7, which agrees with the theorem.

Proof of Theorem 1 Let's imagine that we take away all the edges in the graph. Then the degree of every vertex in the graph is 0. Therefore the sum is also 0. Then we put the edges back into the graph one by one. Subsequently, every time we add an edge to the graph, the degrees of the two endpoints of that edge are increased by 1. Therefore, the sum of the degrees of the vertices is increased by 2. Therefore, the sum of the degrees of the vertices is exactly equal to twice the number of the edges.

Corollary 1. In any graph, the number of vertices with odd degrees is even.

Proof. Let's use the contradiction method (induction) to prove this corollary. Assume there is an odd number of vertices with odd degrees in a graph. Then the sum of the degrees of the vertices is an odd number. But theorem 1 states that the sum of the degrees of the vertices equals twice the number of the edges. Therefore, it's an even number. This is a contradiction.

If G is an undirected graph, a graph consisting of some of G's vertices and edges is defined as a *subgraph* of graph G. For example, in Figure 15.5b, the graph G' is a subgraph of graph G in Figure 15.5a.

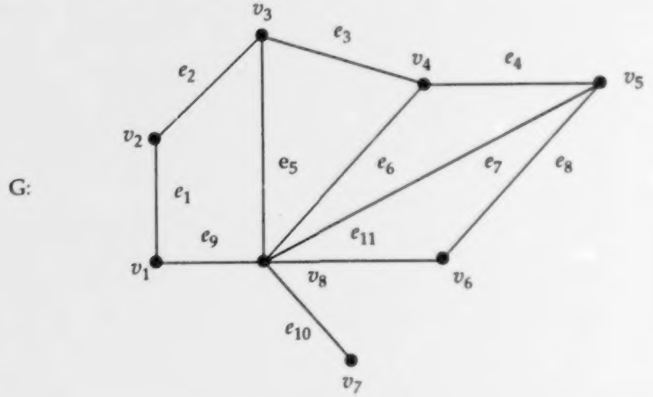


Figure 15.5a

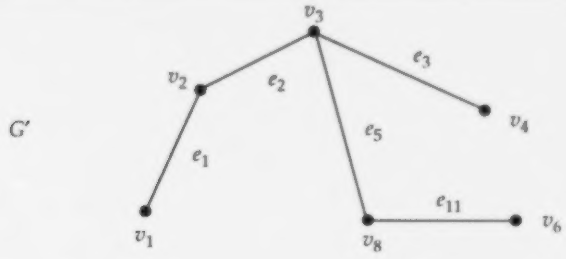


Figure 15.5b

If G is an undirected graph, we define another graph G^c from graph G as follows: G and G^c have the same vertices. For every pair of vertices v_i and v_j in G, if there is an edge joining v_i and v_j in G, then v_i and v_j are not joined in G^c. Conversely, if v_i and v_j are not joined in G, then v_i and v_j are joined in G^c. We call the graph G^c the *complementary graph* of G. For example, the graph in Figure 15.6b is the complementary graph in Figure 15.6a. For the complete graph K_n, its complement is the graph with n vertices and without any edges.

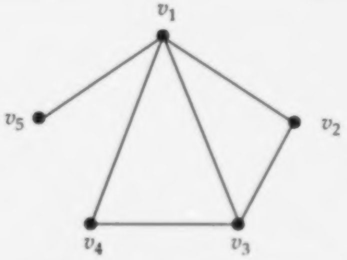


Figure 15.6a

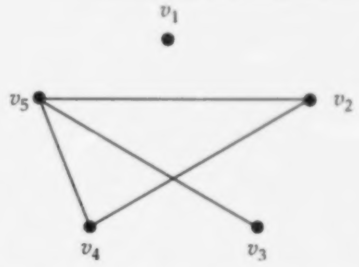


Figure 15.6b

There are many terms in graph theory. We have just listed a few of them here. We will introduce some more when we need them later.

2. Examples

Example 1

> Prove that in any collection of six people either three of them mutually know each other or three of them mutually do not know each other.

This well-known puzzle is a special case of a theorem proved by Ramsey in 1928. The theorem has many deep extensions (see examples 2, 6, and 7) that are important not only in graph theory and combinatorics but also in set theory, logic, and analysis.

This puzzle looks as if it has nothing to do with graph theory. However, let's translate this puzzle into a graph theory problem. Let six vertices v_1, v_2, v_3, v_4, v_5, v_6 represent the six people. If two people know each other, then we use a red edge to join these two vertices. If two people do not know each other, then we use a blue edge to join these two vertices. Since there are edges between every two vertices in the graph, it's a complete graph K_6 with red or blue edges. Now the problem has been translated into the following problem: Use red or blue colors to color the edges in the complete graph K_6. Prove that there must exist either three vertices such that the edges joining them are all red, or three vertices such that the edges joining them are all blue. In other words, there must exist either a red K_3 or a blue K_3.

The puzzle becomes a graph theory problem! Now give the proof the puzzle asks for.

Proof. Let's pick any vertex in K_6, say v_1. The five edges between vertex v_1 and the other five vertices v_2, v_3, v_4, v_5, v_6 are either red or blue. According to the Pigeonhole Principle (see Chapter 8 for detail), at least three edges of the five have the ⑪ same color. Let's assume that $v_1 v_2$, $v_1 v_3$, $v_1 v_4$ are red edges (see Figure 15.7). Now consider the triangle $v_2 v_3 v_4$. If one of the edges $v_2 v_3$, $v_3 v_4$, or $v_2 v_4$ is red, then we will have a red triangle. Otherwise, if $v_2 v_3$, $v_3 v_4$, $v_2 v_4$ are all blue, then the triangle $v_2 v_3 v_4$ is a blue triangle. This proves that there must exist a triangle all of whose edges are colored by the same color.

⑥

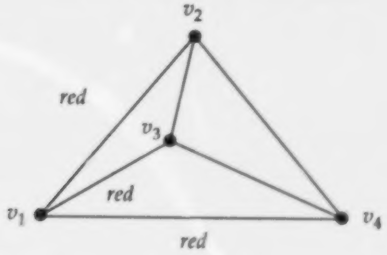


Figure 15.7

There is another way to translate the puzzle into a graph theory problem without using edge coloring. We still use six vertices to represent the six people. If two people know each other, then we connect these two vertices by an edge. If two people do not know each other, then we don't connect these two vertices. For ⑤ example, if everybody knows each other, then it becomes a complete graph K_6. Or if nobody knows each other, then the graph becomes a graph with six vertices without any edges. Of course, those are the extreme cases. Usually there will be some vertices connected with edges and some without. At any rate, the problem we want to solve becomes this:

> Prove that for any undirected graph with six vertices, there must exist either a triangle or three vertices without any edges between them (three isolated vertices).

If we use the terms *subgraph* and *complementary graph*, then the puzzle can be translated to this:

> Prove that in any undirected graph *G* with six vertices, either *G* has a subgraph K_3, or *G*'s complementary graph has a subgraph K_3. (See exercise 1)

Example 2

COMB

> There are 17 scientists who communicate with each other to discuss some problems. In their communications, assume they discuss only three topics. Prove that there are at least 3 scientists who are discussing the same topic.

First, we need to translate this problem into a graph theory problem. Again, we use vertices $v_1, v_2, \ldots, v_{17}$ to represent the 17 scientists. If 2 scientists discuss the first topic, then we use red edges to join these two vertices. Similarly, for the second topic we use blue and for the third topic we use yellow. Then we have a complete graph K_{17} and edges colored red, blue, and yellow. What we need to prove is that there exists a triangle with all edges the same color.

Proof. Choose a vertex, say v_7. The 16 edges from v_7 to the other 16 vertices are colored with red, blue, or yellow. According to the Pigeonhole Principle, at least ⑪ 6 edges have the same color. Let's assume the edges $v_1v_7, v_2v_7, \ldots, v_6v_7$ are colored red. Then let's look at the complete graph K_6 with the vertices $v_1, v_2, \ldots, v_6$. Obviously, if there is a red edge, then we have a red triangle. And if all the edges in K_6 are not red, that means the edges in K_6 are either blue or yellow. Then from the result of example 1, we know that there exists either a blue or a yellow triangle in the graph K_6. That finishes our proof.

Example 2 is a challenging problem. However, using the tools of graph theory, it becomes very easy.

The interesting thing is that if we change the number of people in example 1 from six to five, then we don't have the result. Figure 15.8 is an example. Here, the ⑥ vertices v_1, v_2, v_3, v_4, v_5 represent the five people. The solid edges are colored red, meaning two people know each other, and the dashed edges are colored blue, meaning two people don't know each other. One can't find either a red triangle or a blue triangle in the graph. In other words, no three people know each other, nor do any three people not know each other.

Similarly, if we change the number of scientists in example 2 from 17 to 16, then we don't have the result. We can color the complete graph K_{16} by using three colors (red, blue, and yellow), so that there will be no triangle in the K_{16} with its three edges having the same color. Of course, it will take some time to color the K_{16}. We will leave this for the reader as an exercise.

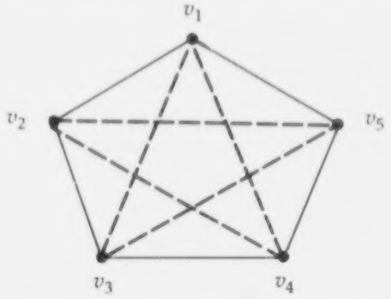


Figure 15.8

Let's look at some more examples related to the degrees of the vertices.

Example 3

COMB

Prove that in any group of people, the number of people knowing an odd number of people (assuming that if A knows B, then B also knows A) is even.

⑥ *Proof.* Since we have the terminology and corollary 1 from section 1, this problem becomes obvious. Let's assume every vertex represents a person. If two people know each other, then we put an edge between these two vertices. The number of people a person knows is just the degree of that vertex. Using corollary 1, we have the solution to this problem.

Example 4

COMB

There are 36 straight lines, colored either red or blue, connecting 9 points on a circle. Suppose that there is at least 1 red line in any triangle formed from the 9 points. Prove that there are 4 points such that every line between 2 of these 4 points is red.

Just by looking at it, you can see this problem is different from examples 1 and 2. But after careful analysis, you will notice the following facts: (a) This problem actually is a graph theory problem. Whether or not the 9 points are on a circle is not important. (b) The graph representing this problem is a complete graph K_9 (the reader should try to find out why this is true). (c) If we can solve the following example 5, then example 4 will be trivial.

Example 5

COMB

Prove: Color all the edges in a complete graph K_9 either red or blue. Then either there are three vertices for which all three edges between these three vertices are blue (a blue K_3) or four vertices such that all six edges between these four vertices are red (a red K_4).

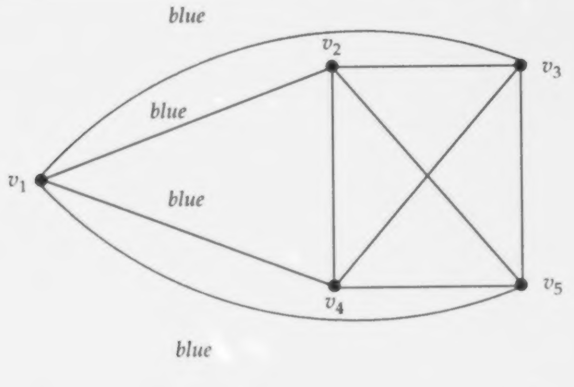
The condition in example 4 is that there is at least one red line in any three-point triangle. Translated to graph theory language, the condition will be, "There doesn't exist a blue triangle." If we proved example 5, then it suggests that there must exist a red K_4 in the example 4. This is exactly the result of the example 4.

Proof. Use $v_1, v_2, \ldots, v_9$ to represent the vertices in K_9. Every vertex is the endpoint ⑪
of eight edges that have been colored either red or blue. Now let's consider several possible cases for the problem:

Case 1. If there exists a vertex, say v_1, such that at least four of the eight edges adjacent to it are colored blue.

Let's fix these four blue edges, say $v_1 v_2, v_1 v_3, v_1 v_4, v_1 v_5$ (see Figure 15.9). Then of the six edges joining the vertices v_2, v_3, v_4, v_5, if one of them were blue then we will have a blue triangle. Otherwise, if all of these six edges are red, then it's a red K_4.

Therefore, the result in example 5 exists in case 1.



⑥

Figure 15.9

Case 2. If there exists a vertex, say v_1, such that of the eight edges joined to it at least six of them are colored red.

Let's fix these six red edges, say $v_1 v_2, v_1 v_3, \ldots, v_1 v_7$. Consider the edges joining the six vertices $v_2, v_3, \ldots, v_7$, and the complete graph K_6 with these six vertices. The edges in that K_6 were colored either red or blue. From the result of example 1, there must exist a triangle with its edges the same color. If there is a blue triangle, then we are done. Otherwise, if we have a red triangle, assume a red triangle $v_2 v_3 v_4$. Since the six edges joining the vertices v_1, v_2, v_3, v_4 are all red, we therefore have a red K_4.

We have proved that we will have the result if case 1 or 2 holds. Now we need to prove that whatever way we color the K_9, at least one of the two cases will ⑬ hold. Actually, if neither case holds, then every vertex has exactly five red edges connected to it, because if there are less than five, we get case 1; and if there are more than five, we get case 2. Now let's look at the subgraph of our K_9 having just its red edges. This graph has nine vertices and every vertex has degree 5. In other words, all nine vertices have odd degrees. But this contradicts corollary 1. This contradiction shows that at least one of the two cases should hold.

The proof in example 5 looks long and complicated. But if you explain the proof by drawing a graph on the blackboard, then you will find it's actually a very easy way to visualize and solve the problem.

The reader should analyze the differences in the proofs of example 5 and examples 1 and 2 and study the way we use the result of corollary 1 in the proof of example 5. That should give a better understanding of using graph theory as a tool to solve problems.

Example 6

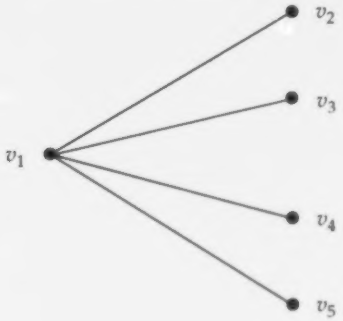
COMB

Nine mathematicians met at an international conference. They found that any two of them have a language in common. If every mathematician speaks at most three languages, prove that at least three of the mathematicians can speak the same language.

Proof. First, we need to construct a graph G representing this situation. The graph G contains nine vertices $v_1, v_2, \ldots, v_9$ representing the nine mathematicians. If two people can speak the same language (no matter what language), then there will be an edge between these two vertices.

Assume that there do not exist three mathematicians who can speak the same language. We will prove that if this assumption exists, then every vertex in the graph G has at most degree 3.

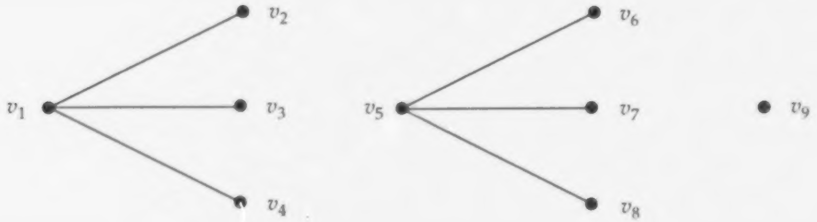
Actually, if there is a vertex of G, say v_1, with degree ≥ 4, then pick four edges ⑪ joined to v_1, say $v_1 v_2$, $v_1 v_3$, $v_1 v_4$, $v_1 v_5$ (see Figure 15.10). Since v_1 can speak at most three languages, then among v_2, v_3, v_4, v_5, there are at least two people who can use the same language to speak to v_1, say v_2 and v_3. But now, v_1, v_2, v_3 can speak the same language. This contradicts the assumption. This proves that the degree of every vertex in G is less than or equal to 3.



⑥

Figure 15.10

Now let's fix v_1. According to what we just proved above, v_1 has at most three edges joined to the other three vertices. Since G has nine vertices, at least $9 - 1 - 3 = 5$ vertices have no edges joined to v_1. Let's assume v_5, v_6, v_7, v_8, v_9 have no edges joined to v_1. Then, similarly, pick v_5 and its degree ≤ 3. Therefore, among v_6, v_7, v_8, v_9, there is at least one vertex not joined to v_5. Let's assume that vertex is v_9. Now we find three vertices v_1, v_5, v_9, and there are no edges between these three vertices (see Figure 15.11). But this contradicts the condition of the problem, which says that there are two people who speak the same language among any three of them. That means that among v_1, v_5, v_9, there is at least one edge between two of the three vertices. From this contradiction, we solve the problem. ⑬



⑥

Figure 15.11

3. Paths, Cycles, Matchings, and Bipartite Graphs

We will use the following graph G (Figure 15.12) to introduce some more important terminology in graph theory.

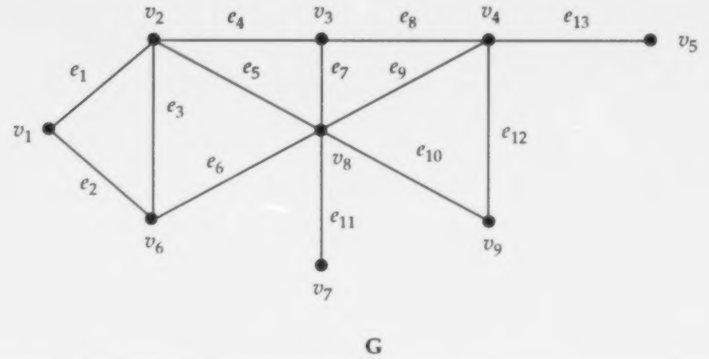

v_2
e_4
v_3
e_8
v_4
e_{13}
v_5
e_1
e_5
e_7
e_9
e_3
v_1
e_{12}
e_6
v_8
e_{10}
e_2
e_{11}
v_6
v_9
v_7
G

Figure 15.12

Consider the sequence of the combination of vertices and edges. For instance, in Figure 15.12, $P_1 = \{v_1, e_1, v_2, e_4, v_3, e_7, v_8, e_{10}, v_9\}$. Here, for every edge, there are two vertices that are the endpoints of that edge. This kind of sequence is called a *path* (see Figure 15.13a). We can say this path starts at v_1 and ends at v_9 or that this path connects v_1 and v_9.

Usually we use a capital letter—for instance P or P_1, P_2, and so on to represent a path.

In a path, vertices and edges may be repeated. For instance, in Figure 15.13b, $P_2 = \{v_2, e_4, v_3, e_7, v_8, e_5, v_2, e_4, v_3, e_8, v_4\}$. Since P_2 satisfies the condition that the vertices before and after every edge are the endpoints of that edge, P_2 is a path. In this path, e_4, v_2, v_3 occurs twice. A path like P_1, having no vertex or edge repeated, is called a *simple path*.

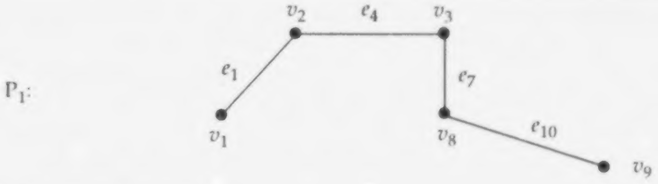

P₁:
v_2
e_4
v_3
e_1
e_7
v_1
v_8
e_{10}
v_9

Figure 15.13a

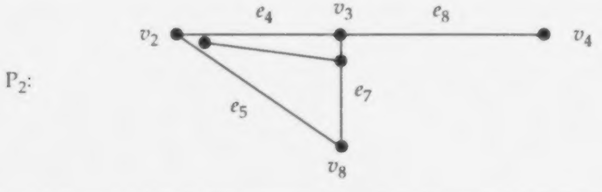
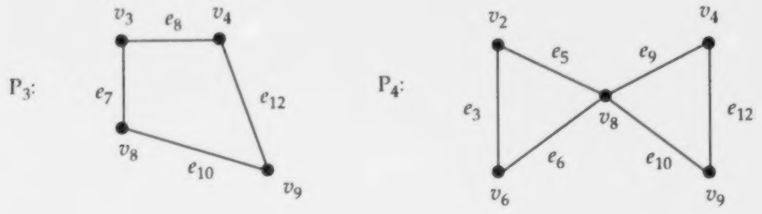

P₂:
e_4
v_3
e_8
v_2
v_4
e_7
e_5
v_8

Figure 15.13b

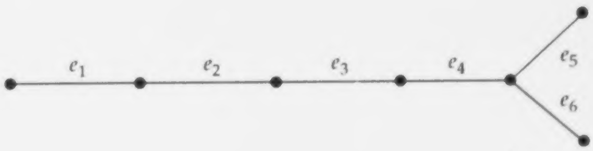
If the first and last vertices of a path are the same, then we call it a *cycle*. Let's use Figure 15.12 as an example: $P_3 = \{v_3, e_8, v_4, e_{12}, v_9, e_{10}, v_8, e_7, v_3\}$ is a cycle (see Figure 15.14a). There are also cycles in Figure 15.14b, such as $P_4 = \{v_2, e_5, v_8, e_9, v_4, e_{12}, v_9, e_{10}, v_8, e_6, v_6, e_3, v_2\}$. Similarly, if no vertices and edges repeat in a cycle (except the first and last vertices), we call it a *simple cycle*.


P3:
v3
e8
v4
e7
e12
v8
e10
v9
P4:
v2
v4
e5
e9
e3
v8
e12
e6
e10
v6
v9

Figures 15.14a and 15.14b

Two distinct vertices or edges in a graph G are *independent* if they are not adjacent in G. A set of pairwise independent edges of G is called a *matching* in G.

For example, in Figure 15.15a, $M_1 = \{e_2, e_5\}$ and $M_2 = \{e_1, e_3, e_6\}$ are matchings in G. Usually we use M, M_1, M_2, and so on to represent the matchings.


e1
e2
e3
e4
e5
e6

Figures 15.15a

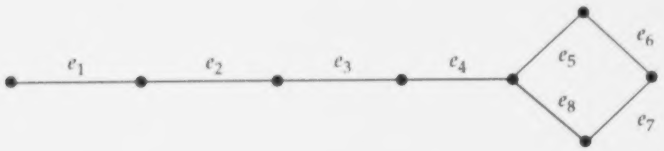

e1
e2
e3
e4
e5
e6
e8
e7

Figure 15.15b

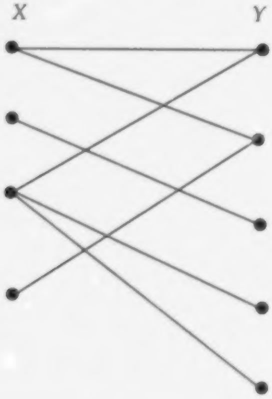
If M is a matching in a graph G with the property that every vertex of G is incident with an edge of M, then M is a *perfect matching* in G.

For example, $M = \{e_1, e_3, e_5, e_7\}$ in Figure 15.15b is a perfect matching in G.

Given a matching M, we usually say the endpoints of the matching are covered by the matching's vertices. Therefore, a perfect matching is one that covers all the vertices of the graph. Since every edge covers two vertices, if a graph has a perfect matching, then it must have an even number of vertices. From that we know that if a graph contains an odd number of vertices, then it doesn't have a perfect matching. For example, Figure 15.15a doesn't have a perfect matching.

If the vertices of a graph can be separated into two parts X and Y so that for every edge in the graph, one of its endpoints belongs to X and the other one belongs to Y, then we call this kind of graph a *bipartite graph.*

Figure 15.16 is an example of a bipartite graph. The left four vertices belong to X and the right five vertices belong to Y. And for every edge in the graph, one endpoint belongs to X and the other one belongs to Y. Therefore it is a bipartite graph.



Figures 15.16

For any graph we can consider its matchings. Matchings in bipartite graphs are very interesting and powerful tools for problem solving.

Theorem 2. If a bipartite graph has a perfect matching, then X and Y must have the same number of vertices.

The proof of theorem 2 is obvious.

Now let's look at some problems related to matchings.

Example 7

> There are some balls. Every ball is colored either red or blue, and every ball weighs either one pound or two pounds. We know that both red and blue balls appear and both one-pound and two-pound balls appear. Prove that there exist two balls with different colors and different weights. COMB

Let's translate the problem into a graph theory problem. We can construct a bipartite graph as follows: X consists of two vertices a_1, a_2, representing the colors red and blue; Y consists of two vertices b_1, b_2, representing the weights one pound and two pounds. If there exists a red one-pound ball, then we put an edge between a_1 and b_1. If there exists a red two-pound ball, then we put an edge between a_1 and b_2, and similarly for the other two cases. Now we have a bipartite graph. The conditions of the problem indicate that (a) both red and blue balls exist, which means that both vertices a_1 and a_2 have edges incident to them, and (b) both one-pound and two-pound balls exist, which means that both vertices b_1 and b_2 have edges incident to them. What we want to prove is that there exist two balls with different colors and weights, which means there exist two edges with different endpoints. In other words, there is a perfect matching. ⑥

Example 7 has now been translated into a graph theory problem as the following: We have a bipartite graph with parts X and Y. Both X and Y contain two vertices, a_1, a_2 and b_1, b_2, respectively. We also know that no vertex has degree 0 (i.e., each vertex has at least one incident edge). Prove that this bipartite graph has a perfect matching.

Proof. Let's pick a vertex from X, say a_1. According to the assumption, there is at least one incident edge, say $a_1 b_1$. Now let's consider a_2; similarly, there is at least one edge incident to a_2. If the edge incident to a_2 is $a_2 b_2$, then $a_1 b_1$ and $a_2 b_2$ will be a perfect matching in the graph, and the problem is solved. If the edge incident to a_2 is $a_2 b_1$ (see Figure 15.17), then consider b_2. As before, there is at least one edge incident to it. If the edge is $a_1 b_2$, then the edges $a_1 b_2$ and $a_2 b_1$ will be a perfect matching in the graph. Otherwise, if the edge at the b_2 is $a_2 b_2$, then $a_1 b_1$ and $a_2 b_2$ will be a perfect matching. Therefore, in all cases, we have a perfect matching. ⑪ ⑬

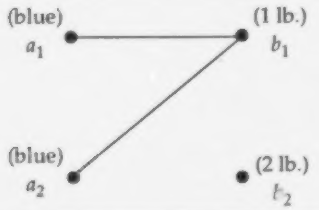


Figure 15.17

Example 8

> COMB There are 10 men and 10 women at a dance. Every man knows exactly 2 women and every woman knows exactly 2 men. Prove that after suitable pairing, every man can dance with a woman he knows.

④ This problem actually is from example 9, but example 9 is much harder. So we will solve this one first as preparation for solving example 9.

Example 8 can be easily translated into a matching problem for a bipartite graph. Let's take 20 vertices, of which 10 represent the men and 10 represent the women. If a man knows a woman, we join the corresponding vertices by an edge. According to the conditions, every vertex has two edges incident to it. That is, every vertex has degree 2 (see Figure 15.18).

We are looking for a perfect matching. So example 8 is exactly the same as the following graph theory problem:

> There is a bipartite graph, the two parts X and Y consisting of 10 vertices each, and every vertex has degree 2. We want to prove that this bipartite graph has a perfect matching.

Proof. Let's randomly pick a_1 to represent a man. There are two edges incident to ⑬ this vertex. Pick any one, say $a_1 b_1$. And from the two edges incident to b_1, pick either one, say $a_2 b_1$. And from a_2, similarly we find $b_2, \ldots$. Since there is only finite number of vertices in the graph, one of the vertices will be repeated eventually. Clearly the repeated vertex will be a_1. Then this is a cycle of the graph. Let's delete this cycle from the graph, and the remaining bipartite graph will still have the property that every vertex has degree 2. Then, using the same method for the remaining graph, we will get another cycle. Repeating this procedure, we can draw the following conclusion: The graph can be split into several cycles. (The graph in Figure 15.19 is an example of this proof.)

It's not hard for us to prove that this bipartite graph has a perfect matching. First, let's notice a fact that every cycle in the bipartite graph has an even number of vertices. From the graph in Figure 15.19 we should see this result (we leave the general proof for the reader as an exercise). Now take any cycle from the bipartite graph. Starting at any edge, we pick the first, third, fifth, ... edges; this is a perfect matching of that bipartite graph. For example, let's first take look at the above cycle in the graph in Figure 15.19. Starting at edge $a_1 b_1$, the second edge is $b_1 a_2$, the third

⑥

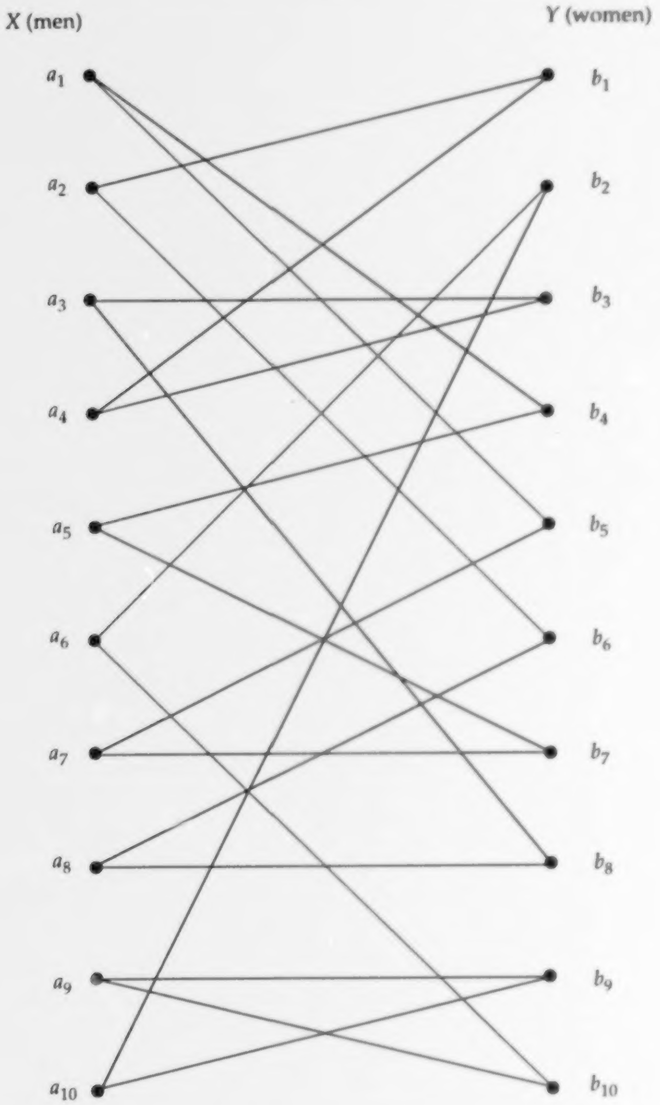

X (men)
Y (women)
a_1
a_2
a_3
a_4
a_5
a_6
a_7
a_8
a_9
a_{10}
b_1
b_2
b_3
b_4
b_5
b_6
b_7
b_8
b_9
b_{10}

Figure 15.18

edge is $a_2 b_2$. Therefore, we pick $a_1 b_1$ and $a_2 b_2$. And now look at the second cycle; we start at edge $b_3' a_3'$, and then we can pick $b_3'a_3'$, $b_2'a_2'$, $b_1'a_1'$. The lines in the graph in Figure 15.19 are these edges.

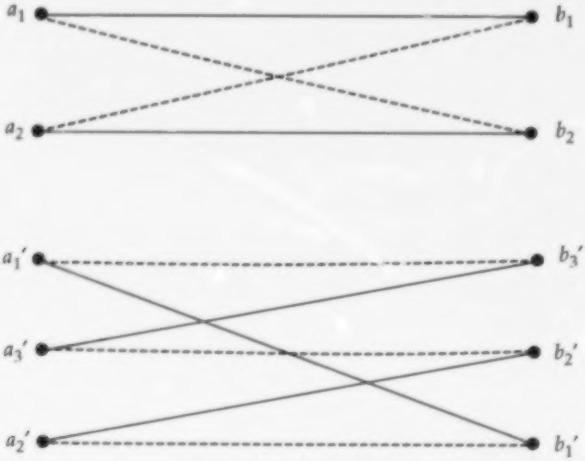


Figure 15.19

The proof of example 8 not only proves the existence of a perfect matching; actually it shows the method of finding a perfect matching in a given graph. In other words, if we have a bipartite graph and all of its vertices have degree 2 (of course, we don't need the condition of each part having exactly 10 vertices; from the proof you should see that the number of vertices can be any finite number), then you just follow this procedure: Pick any vertex and find an edge, then from that edge find another edge, and so on, until we get a cycle. Then continue the above procedure until we find all the cycles of that bipartite graph. When we pick the edges (in odd or even sequences) from each cycle, then we will find the perfect matching(s) in that bipartite graph.

Construct a bipartite graph with its vertices having degree 2 by yourself. Then follow the above procedure to find the perfect matching(s). Doing that by yourself will give you a better understanding of this method.

Example 9

COMB

There are 64 squares in a chessboard. Sixteen of these squares are marked, and there are exactly 2 marks in each row and column. Prove that you can put 8 black and 8 white squares in these 16 marks, and each row and column will have exactly 1 black and 1 white.

Now let's analyze and translate this problem into a graph theory problem. First let's construct a graph (the construction is a little bit special). We consider each row as an vertex. Then we have eight vertices $a_1, a_2, \ldots, a_8$. These eight vertices form the X. Similarly, consider each column as a vertex; then we have $b_1, b_2, \ldots, b_8$. These vertices form the Y. If the i^{th} row and j^{th} column have a mark, then we join a_i and b_j by an edge. This gives us a bipartite graph (e.g., Figure 15.20a is a graph with 16 marks, and the graph in Figure 15.20b is its bipartite graph). Clearly, all vertices in that bipartite graph have degree 2.

After the construction of that bipartite graph, let's see how to use the graph theory language to describe the result that we want to prove for example 9. What we want to prove is that we can put 8 black and 8 white in the 16 marked squares and make each row and column have exactly 1 black and 1 white. Now assume that we already put the 8 black and 8 white in the board as required. Let's see if there is any special property for the 8 black squares. Notice the construction of the graph, in which every square corresponds to an edge in the graph. That is, the squares where 8 black are should correspond to 8 edges. And every row has 1 black, which means that for every vertex in a_i, 1 of the 8 edges is joined to it. In other words, there are no edges with common endpoints. That means these 8 edges are exactly a matching in that bipartite graph. For the same reasons, the 8 white squares and their correspondent 8 edges are also a matching in that bipartite graph. Therefore, what we want to prove in example 9 is the same as the following graph theory problem: If all vertices in a bipartite graph have degree 2, then we can separate them into two groups, and each group is a perfect matching in that bipartite graph.

	B_1	B_2	B_3	B4	B_5	B_6	B_7	B_8
A_1		X				X		
A_2	X				X			
A_3		X					X	
A_4			X					X
A_5	X				X			
A_6			X				X	
A_7				X		X		
A_8				X				X

Figure 15.20a

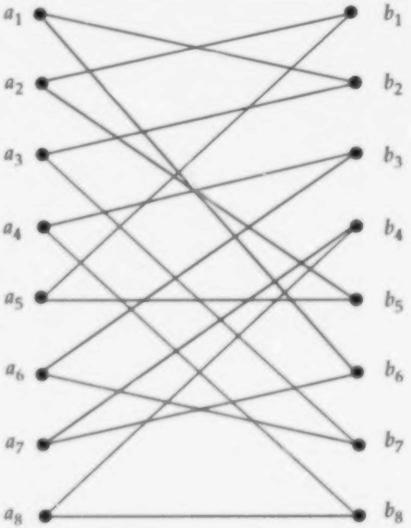


Figure 15.20b

Proof. In the development of the proof in example 8, we saw that if every vertex of a bipartite graph has degree 2, then it can be divided into even cycles. In each cycle, start at any edge, then pick the first, third, fifth, . . . edges that will be a perfect matching in that graph. Obviously, the remaining second, fourth, sixth, . . . edges also yield a perfect matching (see the dashed edges in the graph in Figure 15.19). Since the graph consists entirely of even cycles, this proves that the edges in the graph can be separated into two perfect matchings.

Readers might like to try to solve the next two examples independently, using the method we used to solve example 9. After you do so, you will understand the method of proof much better.

Example 10

COMB

There are *n* pieces of paper. On the front and back of every piece of paper there has been written one of the numbers 1, 2, . . . , *n*, and all numbers appear exactly twice. Prove that we can arrange to put all the pieces of paper on the table, such that the numbers on the pieces of paper show exactly the numbers 1, 2, . . . , *n*.

Example 11

COMB

A mathematics journal requests answers for 10 questions. The answers received by the journal found the following facts: Ten people answered, and every one of them had exactly 2 correct answers, and every question has exactly 2 people who answered it correctly. Prove that the editor can publish the answers of these 10 questions in such a way that all 10 people's names appear and every one of them answers exactly one of the questions.

⑥ Readers should try to translate examples 10 and 11 into graph theory problems. Of course, constructing a correct graph for the problem is very important. Then compare these two problems with example 9 and see if there are any relationships between them.

Examples 10 and 11 are all related to matchings in a bipartite graph. The following is an example of a matching in a non-bipartite graph.

Example 12

> A factory produces two-color products by using six different colors. In the two-color products produced by that factory, every color has been paired with at least three other colors. Prove that we can pick three different products that contain all six colors.

COMB

Let's construct a graph with six vertices *a, b, c, d, e, f,* and every vertex represents one color. If two colors have been used by the factory to produce products, then we join these two vertices by an edge. Now we have a graph. Remember, this is not a bipartite graph.

The condition "every color paired with at least three other colors" means that "every vertex has at least degree 3." And what we want to prove is that we can pick three different products that contain all six colors, which means that there exist three edges, all six of whose endpoints are different. In other words, we want to prove that there exists a three-edge matching. Because this graph only contains six vertices, the problem we actually want to prove is that there exists a perfect matching.

To conclude the above discussion, example 12 can be translated into the following graph theory problem:

> Assume there is a graph with six vertices, and every vertex has at least degree 3. Prove that for this graph there exists a perfect matching.

Proof. Let's consider vertex *a*. Since the degree of vertex $a \geq 3$, there are at least three edges incident to *a*. Let's pick one of them, say *ab*. Then consider vertex *c*; since there are at least three edges incident to *c*, that means that there is at least one edge whose endpoints are not *a* and *b*. Let's say it's *cd*. (See the edge *cd* in the graph in Figure 15.21.)

Our purpose is to find a three-edge matching. Now we already have edges *ab* and *cd*. From there, we will find a matching. For the purpose of clarification: We use the lines to represent the edges of the matching. Clearly, if we have another edge *ef*, then the edges *ab*, *cd*, and *ef* will be a perfect matching in the graph. We solved the problem. So now we need to prove that even if there is no edge *ef*, we will still be able to find a perfect matching in the graph.

Now let's see vertex *e*. We know that there are at least three edges incident to it. Since the endpoint can not be *f*, there are at least three edges from the following edges: *ea*, *eb*, *ec*, *ed*. From there, we know *e* joins at least one of the ⑬ endpoints in both edges *ab* and *cd*. Let's assume *e* is joined to the *a* and *c* (see Figure 15.21, where we use dashes to represent the edges currently not belonging to the matching).

⑥

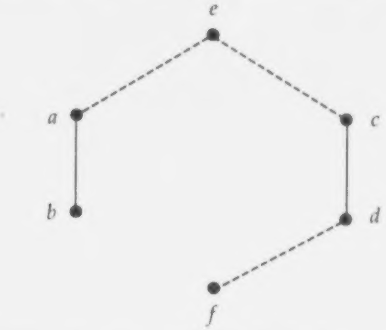


Figure 15.21

Now let's see vertex *f*. Similar to vertex *e*, we assume there is no edge between *e* and *f*. And also vertex *f* has at least three edges incident to it, or *f* has at least three edges from the following edges: *fa*, *fb*, *fc*, *fd*. From there, we know that there is at least one edge from *fb* and *fd*. Let's assume *f* is joined to *d* (see Figure 15.21; of course, there are a lot of edges in the graph—since we don't need them for the proof, we don't draw them). Now, from the edges in the graph in Figure 15.21, it's not hard for us to find a perfect matching in the graph. The method is to take out edge *cd* from the current matching, instead using edges *ce* and *fd*. Then the three edges *ab*, *ce*, *fd* clearly are a perfect matching.

4. Hamilton Paths and Hamilton Cycles

For an undirected graph *G*, if there is a simple path *P* that covers all the vertices in the graph, then we call that path *P* a *Hamilton path*. If a simple cycle covers all the vertices in the graph, then we call that cycle a *Hamilton cycle*.

The thick edges in the graphs in Figure 15.22 are examples of a Hamilton path and Hamilton cycle in the graph, respectively.

Not every undirected graph has Hamilton paths or Hamilton cycles. For instance, the graph in Figure 15.22a does not have a Hamilton cycle. But a lot of interesting mathematical problems are related to the existence of Hamilton paths or Hamilton cycles.

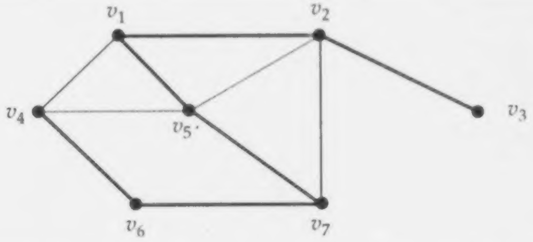


Figure 15.22a

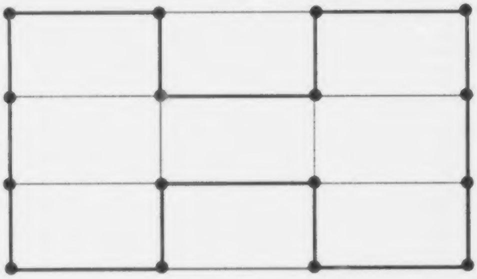


Figure 15.22b

Example 13

In chess, is it possible that a knight can start at the upper-left corner and go through every square on the chessboard exactly once and reach the lower-right corner?

GAME

Let's first look at the movement of a knight in chess. The knight first goes vertically or horizontally for one square and then two squares in a perpendicular direction. Figure 15.23 shows all possible knight moves from square *A*. (See Chapter 18, p. 396)

Now let's translate this problem into a graph theory problem. First, we need to construct a graph. Let's call the squares in the chessboard the vertices in the graph. There are 64 vertices. Let's take a look at two squares; if the knight can move from one of the squares to the other, then we put an edge between these two vertices. For example, in Figure 15.23, the vertices representing the squares *A* and *B* should have an edge joining them. Similarly, there should be edges between *A* and *C*, *A* and *D*,. . . . Then we will have a graph. We are not going to draw this graph.

Now we can translate the problem into a graph theory problem as follows:

Does there exist a Hamilton path from the vertex representing the square *S* to the vertex representing the square *T?*

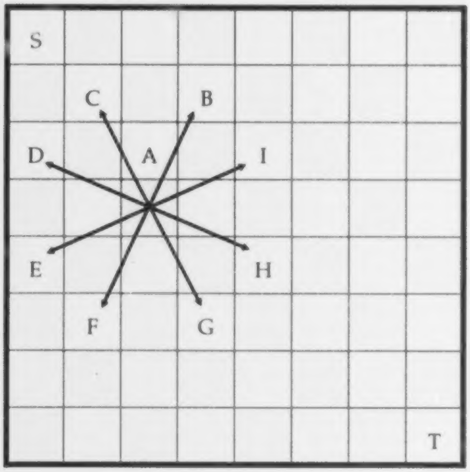


Figure 15.23

Solution. This problem looks very complicated, but since we have the right tools, we will be able to use a very easy method to solve it. First, let's notice that the graph ⑬ in this example is a bipartite graph, because if the knight starts at a white square, it must end at a black square, and vice versa. Therefore, if there is an edge joining two vertices, respective squares must be one white and one black. Clearly, we can separate the vertices into two parts *X* and *Y*. They are white and black squares. And the endpoints of every edge must be such that one of them belongs to *X* and the other one belongs to *Y*. Of course, *X* and *Y* each contain 32 vertices.

In a bipartite graph, if there is a path which has 2, 4, 6, . . . vertices, of the start and end vertices of one of the paths, one belongs to *X* and the other belongs to *Y*, respectively. This fact can be easily explained by the graph in Figure 15.24 (this is a path with eight vertices). If a path's start and end vertices both belong to *X* or both to *Y*, then this path must contain 3, 5, 7, . . . vertices. From that assumption we see that it is impossible for the knight to start at the upper-left corner and then go through every square and reach the lower-right corner. Because if there exists such path, then the bipartite graph will have a Hamilton path with 64 vertices. But the first and last vertices both belong to *X*, since both of them represent white squares.

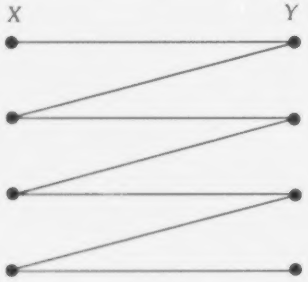


Figure 15.24

Example 14

> There are 10 people at a party. We know that every one of them knows at least 5 of the others. Prove: After suitable arrangement, we can make these 10 people sit around a round table and every one will know the people sitting next to him or her (left and right).

COMB

As in several previous examples, let every vertex represent one person, and there will be an edge joining two vertices if the respective two people know each other. Then we will have an undirected graph with 10 vertices. If we put these 10 people around a round table and let every person know the people sitting next to him or her, this is equivalent to finding a Hamilton cycle in the graph. ⑬

Therefore, the problem can be translated into a graph theory problem as follows:

> There is an undirected graph G with 10 vertices, and every vertex has at least degree 5. Prove that the graph G has a Hamilton cycle.

Proof. Let's use the method of contradiction. Assume that there doesn't exist a Hamilton cycle in graph G. Then we will try to find a contradiction.

Let's consider the two vertices u and v, which are not joined to each other in the graph. Then we add an edge uv to the graph. If, after adding this edge uv, the graph still doesn't have a Hamilton cycle, then we add this edge uv to the graph G. And if, after adding the edge uv, there is a Hamilton cycle, then we stop.

Continue the edge-adding procedure until we don't need to add any more edges. At the end, we will have a graph called G'.

Clearly, the graph G' will have the following properties: (a) G' does not have a Hamilton cycle; (b) every vertex in G' has at least degree 5 (because G' is obtained by adding edges to the graph G, the degrees of vertices in G' are at least as great as in G); (c) after adding an edge in every two unconnected vertices in G', we will have a Hamilton cycle.

Now let's pick two unconnected vertices, say u and v, and add an edge uv as described in (c). We have a Hamilton cycle. When we take out the edge uv, we have a Hamilton path with the endpoints u and v. Let's assume the sequence of the vertices as follows: $v_1\ (= u), v_2, v_3, \ldots, v_{10}\ (= v)$ (see Figure 15.25).

⑥

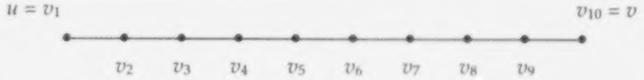


Figure 15.25

Now let's look at vertex v_1. We already know that in G', (a) v_1 and v_2 are connected, (b) v_1 and v_{10} are not connected, (c) the degree of v_1 is at least 5. Therefore, among vertices $\{v_3\}, \{v_4\}, \{v_5\}, \{v_6\}, \{v_7\}, \{v_8\}, \{v_9\}$, there are at least four vertices incident to v_1. We have the same result for vertex v_{10}; there are at least four vertices in $\{v_2, v_3, v_4, v_5, v_6, v_7, v_8\}$ incident to v_{10}.

Now let's look at the seven pairs of vertices: $\{v_2, v_3\}$, $\{v_3, v_4\}$, $\{v_4, v_5\}$, $\{v_5, v_6\}$, $\{v_6, v_7\}$, $\{v_7, v_8\}$, $\{v_8, v_9\}$. We want to prove that among these seven pairs of vertices, there is at least one pair $\{v_j, v_{j+1}\}$ such that v_1 is joining v_{j+1} and v_{10} is joining v_j.

Let's use an example to prove this. As previously mentioned, in $\{v_3, v_4, \ldots, v_9\}$ there are at least four vertices incident to v_1. For instance, let's assume that these four vertices are v_4, v_5, v_8, v_9. Now if there is a vertex in $\{v_3, v_4, v_7, v_8\}$ joining v_{10}, we will have the result. There are also at least four vertices in $\{v_2, v_3, \ldots, v_8\}$ joining v_{10}. So, whatever ways we pick the vertices, there will be at least one vertex from $\{v_3, v_4, v_7, v_8\}$ joining v_{10}.

This proves that there exists a pair of vertices $\{v_j, v_{j+1}\}$ such that v_1 joins v_{j+1} and v_{10} joins v_j. But now there exists a Hamilton cycle in the graph G'. That is $v_1, v_2, \ldots, v_j, v_{10}, v_9, \ldots, v_{j+1}, v_1$ (see graph in Figure 15.26, for which this is the example when $j = 4$). This contradicts our assumption that there is no Hamilton cycle in G'.

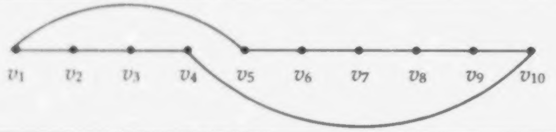


Figure 15.26

If we change the 10 people to 2^n people, and the degree of every vertex to n, the result of this example still exists (see the exercises).

5. Tournaments

This section introduces a set of special directed graphs, in which there is one arc between every two vertices in the directed graph. We call these graphs *tournaments*. The two graphs in Figure 15.27 are two examples of tournaments with 4 and 5 vertices.

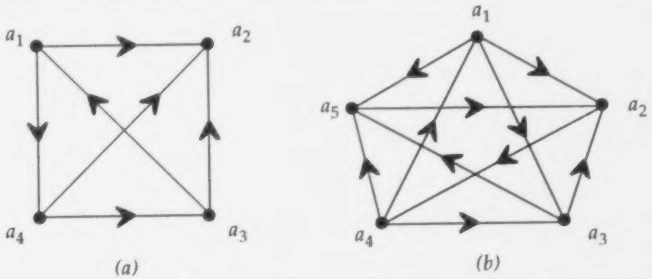


Figure 15.27

Why do we call these tournaments? Because such a graph can represent the result of a round-robin tournament. For instance, in the graph in Figure 15.27b, we assume every vertex a_i represents a team, and an arc from a_i to a_j means a_i beats a_j (assume there is no tie). Therefore, the graph represents the results of a tournament.

Now let's introduce some more terminology related to directed graphs.

First, in a directed graph, we can also consider paths and cycles. If the arcs in a path or cycle all have the same direction (see graphs in Figure 15.28), then we call them a *directed path* or a *directed cycle*.

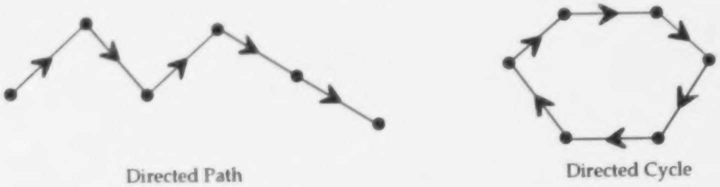


Figure 15.28

Similarly, if a simple directed path or a simple directed cycle goes through all the vertices in the graph, then we call it a *directed Hamilton path* or a *directed Hamilton cycle.*

If $a = (u, v)$ is an arc of a directed graph G, then a is said to join u to v. We further say that a is incident from u and incident to v, while u is incident to a and v is incident from a. Moreover, u is said to be adjacent to v and v is adjacent from u. In the directed graph in Figure 15.27a, vertex a_1 is adjacent to vertex a_2, but a_2 is not adjacent to a_1. Two vertices u and v of a directed graph G are nonadjacent if u is neither adjacent to nor adjacent from v in G.

The *outdegree* of a vertex v, or od (v), of a directed graph G is the number of vertices of G that are adjacent from v. The *indegree* of v, or id (v), is the number of vertices of G adjacent to v. The *degree* of a vertex v, or deg (v), of G is defined by the sum od (v) + id (v). In Figure 15.27a, od $(a_1) = 3$, od $(a_2) = 0$, od $(a_3) = 1$, od $(a_4) = 2$, id $(a_1) = 0$, id $(a_2) = 3$, id $(a_3) = 2$, id $(a_4) = 1$.

Example 15

GAME

> There are n ($n \geq 3$) teams attending a tournament. Assume there is no tie in the game. Prove that, whatever the results of each game, we will be able to arrange these n teams into a sequence a_1', a_2', . . . , a_n' such that a_1' beats a_2', a_2' beats a_3', . . . , and a_{n-1}' beats a_n'.

Clearly, we can translate this problem into a graph theory problem:

Prove that in any tournament there exists a directed Hamilton path. Let's take a look at the examples in Figure 15.27; the two graphs in Figure 15.29 are Hamilton paths.

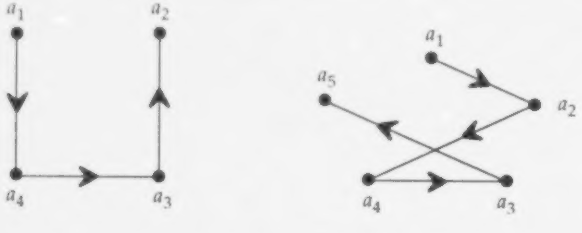
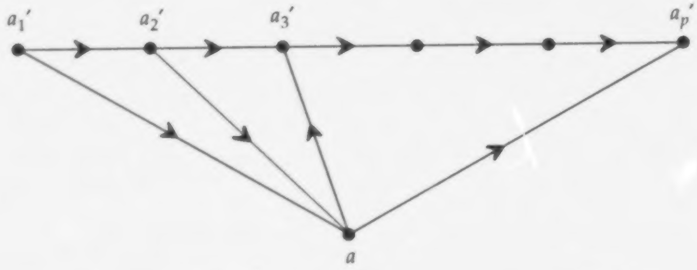


Figure 15.29

Proof. The proof procedure actually is a method of finding the required Hamilton path. Let's first pick any directed path $a_1', a_2', \ldots, a_p'$. We will prove that if this path does not contain all the vertices in the graph, then we can extend this path. For example, say a is not in the path. Now let's look at the following: (a) If the arc ⑬ connecting the vertices a and a_1' is from a, then $a, a_1', \ldots, a_p'$ is a longer directed path; (b) similarly, if the arc connecting vertices a and a_p' is from a_p', then $a_1', a_2', \ldots, a_p', a$ is a longer directed path; (c) if both (a) and (b) don't exist, then there must exist the arcs from a_1' to a and a to a_p' (see Figure 15.30).



⑥

Figure 15.30

Now let's observe the following sequence $a_2', a_3', \ldots, a_p'$, where a_j' is the first vertex such that an arc goes from a to a_j' (see a_3' in Figure 15.30). Then clearly the ⑥ path $a_1', \ldots, a_{j-1}', a, a_j', \ldots, a_p'$ is a longer directed path.

Therefore, if a directed path does not contain all the vertices in the graph, then we can always extend this path. This proves that there must exist a directed Hamilton path, that is, one that contains all the vertices in the graph.

Example 16

There are n people in a tournament. Prove that, after the tournament, there is at least one person who can announce, "For all of you, either I beat you or you lost to someone I beat." GAME

Let's use graph theory language to describe the problem:

Prove that in a tournament with the vertices $a_1, a_2, \ldots, a_n$, there must exist a vertex a_i such that for every other vertex a_j, either there is an arc from a_i to a_j or there is a directed path from a_i to a_j with two arcs (if this path's vertices are a_i, a_k, a_j, then a_i beats a_k and a_k beats a_j. Therefore, a_j lost to a_k and a_i beats a_k).

Proof. We want to find a person who wins the most (The answer may be not unique—why?), that is, the vertex that has the largest outdegree. Let's assume this vertex is a_1 (see Figure 15.31), and assume a_1 beats $a_2, a_3, \ldots, a_s$, that is, a_1 wins s – 1 times. But a_1 lost to $a_{s+1}, \ldots, a_n$. What we want to prove is that for every vertex a_j in $\{a_{s+1}, a_{s+2}, \ldots, a_n\}$, there must exist an vertex a_k in $\{a_2, a_3, \ldots, a_s\}$ such that a_k beats a_j; that is, there is an arc from a_k to a_j.

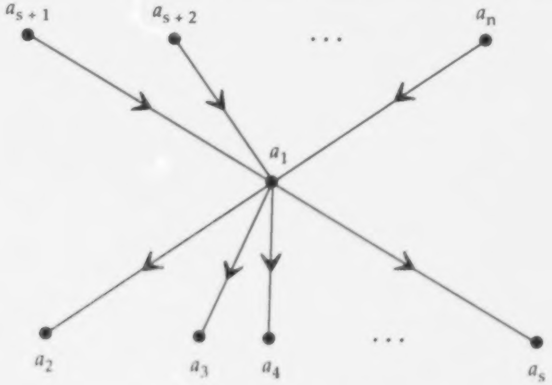


Figure 15.31

Let a_j be any vertex in $\{a_{s+1}, a_{s+2}, \ldots, a_n\}$, so a_j beats a_1. Now let's consider the result between a_j and $\{a_2, a_3, \ldots, a_s\}$. If a_j beat all of these teams, then a_j won s times. Therefore, a_j wins more times then a_1. This contradicts the assumption of the problem: a_1 wins the most. This proves that in $\{a_2, a_3, \ldots, a_s\}$, there must exist an a_k such that a_k beats a_j.

(13)

Example 17

GAME

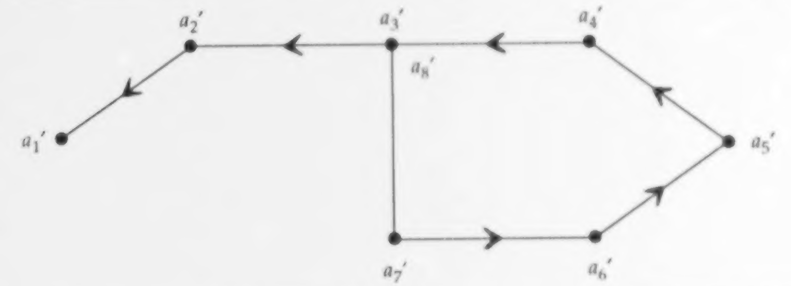
There are n ($n \geq 3$) teams $a_1, a_2, \ldots, a_n$ in a tournament. Assume that there is no team that always won. Prove that there must exist three teams a_i, a_j, a_k such that a_i beats a_j, a_j beats a_k, and a_k beats a_i.

Let's use graph theory language to describe the problem we want to prove: There is a tournament that satisfies the following condition: Every vertex's indegree is not equal to zero (i.e., every team lost at least once). Then there exists a directed cycle with three arcs.

Proof. We need two steps to prove this fact.

First, we will prove that there must exist a directed cycle (more than three arcs allowed) in the assumed tournament.

Pick any vertex in the graph, say a_1'. According to the assumption, a_1' has indegree at least one. Assume this arc starts at a_2'. Similarly, there is at least one arc to a_2'. Let's assume that arc start at a_3'. Following this procedure, we will get $a_1', a_2', a_3', \ldots$. Since there is only a finite number of vertices in the graph, some ⑬ vertex will be repeated eventually. Let's assume that the first repeated vertex is a_s', which is the same as some previous vertex a_t'. Clearly, we have found a directed cycle (see graph in Figure 15.32, where a_8' and a_3' repeated, and we have a directed cycle $a_8'\, a_7' \ldots a_4'\, a_3'$).



⑥

Figure 15.32

Second, we now need to prove that in a tournament, if there is a directed cycle with more than three arcs, then we will always be able to find a directed cycle with a lesser number of arcs. Let's take the graph in Figure 15.33, which has a directed cycle with seven arcs, as an example. Let's consider the arc between a_1 and a_3. If the direction of that arc is from a_3 to a_1, then we will have an directed cycle with three arcs: $a_1\, a_2\, a_3\, a_1$. If the direction of that arc is from a_1 to a_3, then we will have a directed cycle with six arcs: $a_1\, a_3\, a_4\, a_5\, a_6\, a_7\, a_1$. Therefore, in any case, we will have a directed cycle with a lesser number of arcs. This proves that there must exist a directed cycle with three arcs.

⑥

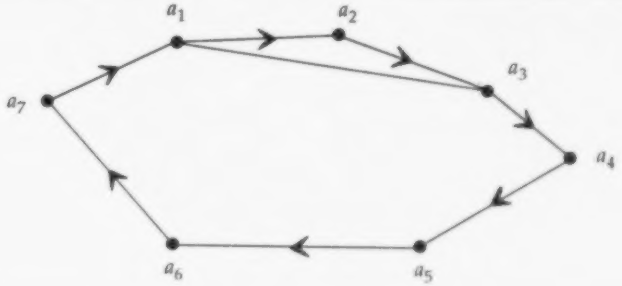


Figure 15.33

6. Exercises

The following are some exercises for readers to practice using graph theory as a tool to solve the problems. The starred (*) problems are a little bit difficult, but readers are encouraged to read the similar examples in the chapter and try to solve them.

1. Prove: For any undirected graph G with six vertices, either G has K_3 as subgraph, or the complementary graph of G has a subgraph K_3.
2. A club has 99 members. Everyone in the club knows at least 66 other members. Prove that we can pick 4 of them to play bridge such that all 4 know each other.
3. In exercise 2, if every member of that club knows exactly 66 of the 99 club members, prove that we may not be able to pick 4 of them to play bridge such that all of them know each other.
4. There are 9 people $v_1, v_2, \ldots, v_9$. Suppose that v_1 shook hands with 2 people; v_2, v_3 shook hands with 4 people; v_4, v_5, v_6, v_7 shook hands with 5 people; and v_8, v_9 shook hands with 6 people. Prove that we can find 3 of these 9 people such that all 3 of them shook hands with each other.
5. The edges in the complete graph K_{18} are colored either red or blue. Prove that whatever way we color the edges, there must exist four vertices such that all six edges between these four vertices (a K_4) are colored by the same color.
6. Prove: Among 14 people, either 5 of them mutually know each other or 3 of them mutually do not know each other.

*7. Prove: Among 19 people, either 6 of them mutually know each other or 3 of them mutually do not know each other.

8. We color the edges in the complete graph K_{66} by using four colors: red, blue, yellow, green. Prove that whatever way we color the edges, there

exists a triangle and all three of its edges (a K_3) are colored by the same color.

*9. Prove: If there are n ($n \geq 2$) people in a meeting, there must exist 2 people who know exactly the same number of people in the meeting.

10. Prove that the cycle(s) in a bipartite graph must contain an even number of vertices and an even number of edges.

11. There are n men and n women at a dance. We already know that each man has danced with at least 1 woman but has not danced with all the women. Similarly, we know that each woman has danced with at least 1 man but has not danced with all the men. Prove: There exist 2 men x_1', x_2' and 2 women y_1', y_2' such that the following pairs have danced before: x_1', y_1' and x_2', y_2'; but the following pairs have not danced before: x_1', y_2' and x_2', y_1'.

12. Find a perfect matching in the bipartite graph in Figure 15.18.

13. There are 6 men and 6 women attending a dance. After the dance, they recall the number of people they have danced with. The numbers are as follows: 3, 3, 3, 3, 3, 5, 6, 6, 6, 6, 6, 6. Prove that at least one of them made mistake.

14. Try to put 8 white and 8 black chess pieces in the squares with marks "X" in Figure 15.20a, and make every row and column have exactly 1 white and 1 black.

15. There is a classroom with 25 seats arranged as a square of 5 rows and 5 columns. Assume that at the beginning every seat has a student sitting in it. Can we change the students' seats in such a way that every student can be moved to all adjacent seats?

16. Construct a graph which has 40 vertices and 400 edges but no triangles.

*17. Prove that any graph with 40 vertices and 401 edges must contain triangle(s).

*18. Prove: If a graph G has $2n + 1$ vertices and $n^2 + n + 1$ edges, then the graph G must contain triangle(s).

*19. There are $2n$ people at a party. We already know that everyone knows at least n people. Prove that we can pick 4 people from these $2n$ people such that these 4 people sit at a round table and every one of them knows the people sitting on either side of him or her.

*20. There are $2n$ people at a party. We already know that each person knows at least n people. Prove that after suitable arrangement, we can have these $2n$ people sit around a round table and every one of them knows the people sitting on either side of him or her.

21. Decide if there is a Hamilton path or Hamilton cycle in the grid graph in Figure 15.34. (Hint: Consider the two cases in which n is odd or even.)

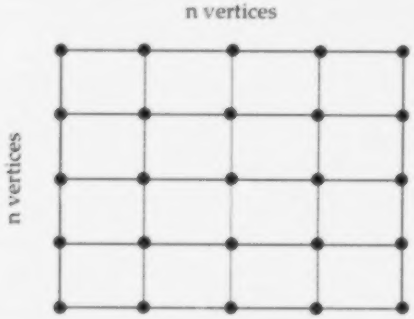

n vertices
n vertices

Figure 15.34

22. Prove that there is no Hamilton cycle in the graph in Figure 15.35.

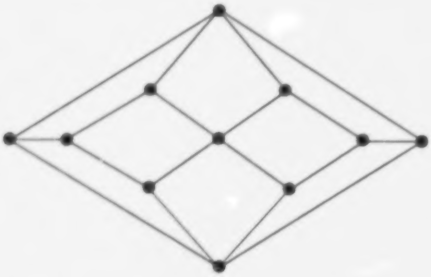


Figure 15.35

16

A Different Solution for Problems With Extreme Values

A Didactic Journey Into the World of Jakob Steiner's Ideas

KARL KIESSWETTER

ROLAND J. K. STOWASSER

LENNI I. HAAPASALO

Swiss mathematician Jakob Steiner (1796-1863) did most of his mathematical work in Berlin. He was especially interested in geometry, and the solution of the isoperimetrical problem is his best-known achievement. The approaches to this and similar problems are fascinatingly elegant and simple, easily lending themselves to use in teaching. In this chapter, we will write and rewrite selected parts of his work on the topic mentioned and prepare it for direct use in the classroom. It is significant that the great Karl Weierstrass not only corrected and completed Steiner's proofs but also later edited the complete work of Jakob Steiner. By 1841, Steiner had already addressed our main concern:

> Of the two approaches for the investigation of features which are responsible in geometric figures for maximum and minimum, the synthetic one has been very neglected and it is believed that it is an inadequate approach which has to be replaced. It is thought that only the other approach, the analytic one, has all the advantages. However, the common directions, which are available in the analysis for this special purpose, don't easily lead us in many cases to the aim. Often, they don't seem to be appropriate for revealing the real being or true reason for the maximum and minimum, and in many cases they only show features closely or more distantly related to the cause, but not the cause itself. That is the reason it seems to be more useful to select another viewpoint or, better still, to return to the abandoned method to search for all real causes which are the reasons for the maximum and minimum, so there are different fundamental features from which we can derive a system of closely interrelated theorems.

In order to show that Steiner (1971) didn't think only geometrically, we offer another citation:

> Although I may now think that the synthetic method is the most appropriate one for observation and reasoning in these fundamental theorems and their future development, it is possible that the analytic approach could be the correct viewpoint in later forthcoming questions to pursue the subject in appropriate cases.

How is teaching done in school today? Problems with extreme values are not approached from one-sided viewpoints; that is, they are not restricted to the best-known analytic models. Rather, it is the aim in mathematics and, accordingly, in the teaching of mathematics, to develop comprehensive exemplary solutions. But in the teaching of mathematics, one must not suggest that such solutions are unique and that it is sufficient merely to find them. Above all, it should not be the aim to make repeated use of only one method, like a cooking recipe. In addition, methods other than Steiner's (polynomial transformation, contour lines, recursions, and so forth) should be a permanent part of the teaching process.

Steiner focuses his endeavors on the main theorem: that, as he observed, the circle has the greatest area of all plane figures with a given constant boundary. As is typical for him, he proves this theorem in several different ways. He traces his procedure of symmetrization to the following "first fundamental theorem": *Of all triangles having the same base and the same perimeter, the isosceles triangle has the largest area.*

Another proof, which he used in later works and which is very simple, is based on Steiner's "second fundamental theorem": *If two sides of a triangle are given, then the area of the triangle is largest if the two sides include a right angle.*

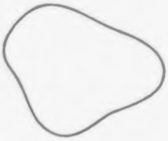
These ideas for proofs readily present themselves for teaching—particularly the second fundamental theorem, since it is much more easily proved by students than the first.

In teaching lessons, one needs to focus very intensively on the initial problem. Of what use is it? It serves as a motivation. Its formulation must not create difficulties because of too many new and unfamiliar elements. Connections should be made with circumstances that are familiar and as concrete as possible. One has to create an open situation, thus guaranteeing, at the same time, that the students get ideas that lead to a solution. The formulation of the initial problem should lead the students to make appropriate associations. They must be able to use familiar or easily found heuristic strategies, and the complexity of the solution should be at a level sufficient to provide a challenge and encourage efficiency. However, it is possible for the complexity to be at too high a level, because too many reductions and other auxiliary problems that arise in the course of arriving at a solution may create confusion, and students may tire. The student should draw on his or her strengths and should recognize that there is a great need for imagination (not tricks) and that, in order to solve problems, one must continue to try. With this viewpoint in mind, we want to discuss the initial problem. Here are some proposals.

Problem 1

Given a closed rope with length *l*. How should the rope be placed so that it borders the largest possible area? What size is this area?

GEOM



⑥

Figure 16.1

The formulation of the problem is simple and doesn't contain unfamiliar aspects. Also, particularly in the optical field, there is no direct connection to the circumstances described in the second fundamental theorem of Steiner. Furthermore, the number of points where the position of the rope can be changed has the potency of the continuum and so cannot be readily grasped.

Problem 2

GEOM

Given a simple closed set of connected edges in a plane, with the lengths of the edges $s_1, s_2, \ldots, s_n$. Imagine that the edges are bars from a construction set that are connected with joints. How should this set of edges be placed so as to border the largest possible area?

⑥

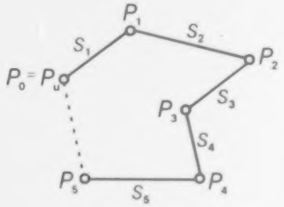


Figure 16.2

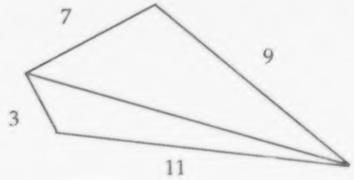
The formulation of the problem contains a connection to familiar situations in life as well as a mathematically formulated text. In the optical field, one can draw an extensive relation to the second fundamental theorem. The advantages of this formulation is that changes can be made only on a finite number of points (joints, bars). However, a disadvantage is that the solution requires a sequence of several steps, and students often need help. There is the risk that the teacher may have to provide help too often, and students may feel that the teacher is using tricks at different stages of the work.

Problem 3

(A contrived story!) About 2,500 years ago, if someone wanted to join the Pythagoreans, he had to demonstrate his talents. Pythagoras himself is said to have proposed the following problem for this purpose (for simplification, sides as well as lengths of sides are designated the same way):

GEOM

Given a convex quadrilateral with sides $s_1 = 7$, $s_2 = 9$, $s_3 = 3$, and $s_4 = 11$. What should be the measure of the angle between s_1 and s_2 in order for the quadrilateral to have the largest possible area?



⑥
⑧

Figure 16.3

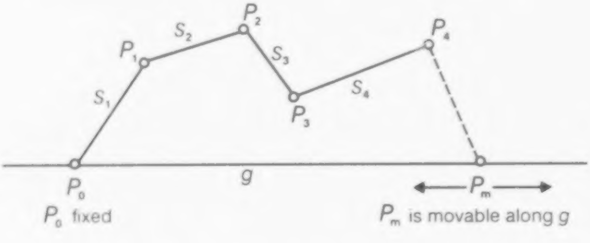
This formulation of the problem will result in direct and strong associations with the Pythagorean theorem and then to the rectangle. The quite independent discovery of the second fundamental theorem should now be only a question of time. One should easily recognize that $7^2 + 9^2 = 11^2 + 3^2$ and that this is a sufficient condition for a quadrilateral to have maximum area. Despite this, it is not a satisfying solution from a didactic viewpoint, because there are almost no possibilities for generalization; the route to making general connections is blocked.

Problem 4

5

Given a simple open set of connected edges with the lengths of the edges $s_1, s_2, \ldots, s_m$, which lies on one side of a straight line g, between the end points P_o and P_m (on g) (conditions as in problem 2). How should this set of edges be placed so that, along with the edge $\overline{P_o P_m}$, the largest possible area is enclosed?

GEOM



⑥

Figure 16.4

Of course, this isn't a general optimal formulation, but we think it's the best compromise. Its advantages are that students are able, with comparative ease, to find their own ideas for the solution and that the entire solution, as well as individual elements of it, can be used for similar problems in an appropriate way. Along with the elements of the proof, the student gets an instrument for a relatively elegant and transparent solution for a whole class of problems with extreme values. We want to illustrate a detailed approach to the solution of problem 4. The reader can formulate his or her own strategies.

Approach and Elements of Solution

Step 1

A common heuristic strategy (we hope!) is to observe special cases. For $m = 2$ we have Figure 16.5.

⑤

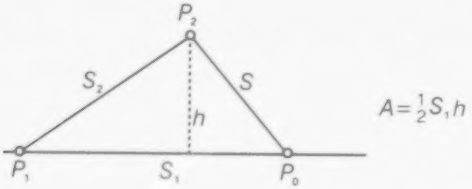


Figure 16.5

Every student knows this formula for the area. But he or she must also realize that movement of P_2 straight toward P_0 causes oppositely oriented changes in g and h. In addition, it is not self-evident that the student can effect a change to a simpler form by taking another viewpoint, using a rotation.

③

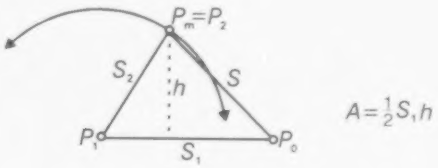


Figure 16.6

Now one need only address one changing parameter (*h* in Figure 16.6), whose maximum value is obviously s_2 The area *F* is maximized only when the angle between s_1 and s_2 is a right angle.

As we mentioned, this is such an important fact in Steiner's work that he speaks of it as a fundamental theorem.

Step 2

It is easy to conclude that none of the possible regions could have maximal area if it is not convex (turn $\Delta P_2P_3P_4$ outwardly in Figure 16.4).

Step 3

Let's have a look at Figure 16.7. Consider the hatched regions over $\overline{P_o\,P_k}$ and $\overline{P_k\,P_m}$ as fixed, and shift the final point P_m along *g*. Only the drawn angle α_k at P_k changes. But in the meantime, the area of $\Delta P_oP_kP_m$ changes as the angle changes. The area is maximal if α_k is a right angle.

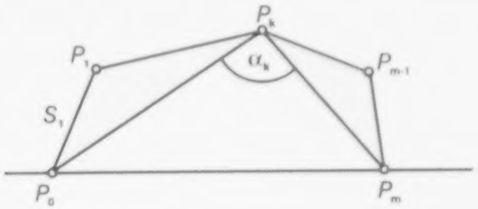


Figure 16.7

What can we learn?

1. The area of the surface cannot be maximal if one of the angles α_k, $k = 1, 2, \ldots, (m - 1)$ is not a right angle.
2. If one of the angles α_k, $k = 1, 2, \ldots, (m - 1)$ is not a right angle, then the surface can be enlarged by an allowed change at the border line.
3. It is not possible to enlarge the surface in an appropriate way by this procedure if all angles α_k, $k = 1, 2, \ldots, (m - 1)$ are right angles; that is the case when all points P_k, $k = 1, 2, \ldots, m$ lie on a semicircle with diameter $\overline{P_o\,P_m}$.

Step 4

But that is not all. Many students are going to find learning these concepts difficult. The level of success can be even greater if other thoughts are added to complete the proof (from another part of mathematics, namely, analysis). One remaining question is still unanswered: assuming that border lines such as those in problem 4 are given, is it always possible to inscribe them in a semicircle in the desired way?

To answer that question, considerations of continuity (intermediate value concept) must be made. Those could be replaced by the following consideration of plausibility (whose positive pedagogical value should not be underestimated).

We proceed from the semicircle having a diameter $d = s_1 + s_2 + \ldots + s_m$. We can lay off the edge drawing on this semicircle, as shown in Figure 16.8, because the semicircle arc is longer than its diameter. Then we steadily reduce the size of the semicircle and let the "final point P_m wander into the point E." This is always possible, because we can find semicircles that are surely too small to lay off the edge drawing in the manner described (e.g., $d = \max \{s_k \mid k = 1, 2, \ldots, m\}$).

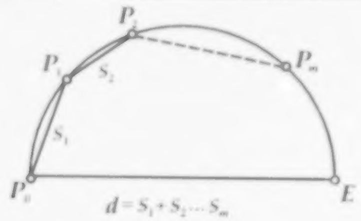


Figure 16.8

Step 5

To emphasize which final gap must be closed, we'll have a closer look at the third step. There we chose a procedure that enabled us to specify in each case a larger surface from among those that were under consideration, aside from the "semicircle case." But perhaps our procedure is useless for precisely this semicircle case. Perhaps another procedure also gives us an enlargement of surface area. A good way to get a feeling for this problem is through the following example.

Let's consider the open interval $I := (0, 2)$ and the "enlargement instruction" $x \rightarrow f(x) = x + (2 - x)(x - 1)^2$. Then $f(x) \in I$ for all $x \in I$, and $x < f(x)$ for all $x \in I$ with $x \neq 1$, but $1 = f(1)$.

Thus, we have an analogous fact. Obviously from this we cannot conclude that 1 is the maximum of $f(x)$ in I.

But should it be possible to prove the existence of a maximum, then we must have a maximum in our semicircle case, because that is the only case for which the improving procedure (enlarging the area) does not work.

If it is clear what is missing, the teacher can easily proceed without the difficult proof. However, he or she can point out that a great mathematician like Steiner himself left a gap at this point. Finally, we would like to show that the manner in which we set up problem 4 is so helpful that even this missing proof could be understood by good students. It is to our advantage that we are able to describe the course of our border line with a finite number of variables.

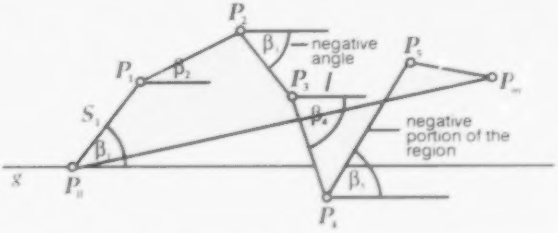


Figure 16.9

A preliminary consideration leads to the determination that we need to consider only those border lines in the figure shown where "P_k lies to the right of P_i" when k is greater than i. In our proof, for technical reasons, we eliminate the requirement that the final point P_m lies on g, thereby enlarging the number of covered border lines again. Our only condition is that, for all angles β_k of the individual edges with the straight line g, the inequality $-90° \leq \beta_k \leq 90°$ is valid. Then each of those border lines is represented by a vector out of the closed n-dimensional cube: $W_m = [-90°, 90°]^m = \{(x_1, \ldots, x_m) \mid -90° \leq x_k \leq 90°\}$.

As indicated in Figure 16.9, each of those vectors out of W_m is attached to the surface area "under that border line" that it represents. As a result, surface segments and the line segment $\overline{P_o P_m}$ are rated negatively. It is "obvious" that we obtain a continuous function. That is all for the mathematician, because the continuous image of a compact set is itself also compact. All we need in order to solve our problem is to know that the image of a closed cube has a maximum under a continuous function, which can be shown in an analogous special case $m = 1$, which is usually treated at school.

Additional Considerations and Problems

The solution of our problem 4 creates an open situation, which should be used in accordance with Polya's ideas. One is likely to ask which similar problems emerge and how they can be solved. One example is problem 2. If we use the "method of fixed-edge surfaces," as we did in the third step mentioned above, we can trace back our already solved problem. However, it is more useful to treat the following problem directly.

Problem 5

GEOM Given an open rope, how should it be placed so that, together with the segment $\overline{AE}$ connecting its endpoints, it encloses a maximal surface?

⑥

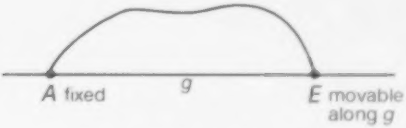


Figure 16.10

④ ⑤ To show that the maximal surface occurs with a semicircle, students can work analogously to the solution of problem 4 (transfer, consolidation). Only the transformation of thoughts from step 5 leads to greater difficulties. The teacher should be satisfied with plausibility considerations, noting that the circle is the limiting case of a suitably chosen regular polygon of *n* vertices (equal length).

Problem 1 then easily leads back to problem 5. The given closed rope is now divided into two equal open ropes. With that, the isoperimetric main clause is proved. With the help of the method of fixed-edge surfaces and other things, this can be used to obtain another solution for problem 2. We give still another example to illustrate this approach.

Problem 6

GEOM The ends of a rope are fastened also to the endpoints of bar *S*. How is the rope to be placed so that the area of the enclosed surface is maximized?

⑥

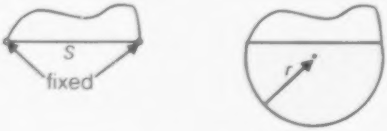


Figure 16.11

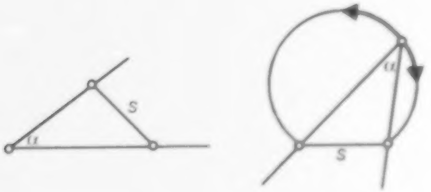
Solution

There is exactly one circular arc region smoothly bounded by the rope and the bar. Its radius is r. On the other side of S the rest of the circular area is to be completed (hatched in Figure 16.11). The entire area is then maximal only if the region is circular. Therefore, . . . there are several extreme-value problems of the kind mentioned above. In approaching these problems, one cannot always follow the detailed solution elements discussed. However, reflecting upon the simplest kinds of elementary concepts is normally sufficient (e.g., problem 7).

Problem 7

Given an angle α and a bar S, so that S links the two sides of α. In which position should S be placed so that the largest possible surface is enclosed?

GEOM



⑥
③

Figure 16.12

Another point of view and restructuring of the situation brings (as in the first step) the realization that the isosceles triangle has the maximum area (Figure 16.12). One realizes that problems 6 and 7 could both be used to solve an additional, similar problem, as follows.

Problem 8

GEOM

Given the angle α and a rope of length *l*, whose endpoints *A* and *E* can be shifted along the sides of α. How should one position the rope so that the largest possible surface is enclosed?

⑥

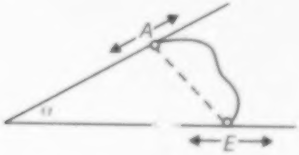


Figure 16.13

(If the reader does not have the solution, he or she should recall the method of fixed-edge surfaces.)

Epilogue

We hope that the reader had fun with Steiner's "World of Ideas" and that this world of ideas will bring pleasure to students as well.

Our objective will have been achieved if the reader uses some of the problems given here in his or her lessons and thereby improves the spectrum of the mathematical handling of extreme-value exercises. Because Steiner's works are good reading, one can glean some valuable suggestions, such as the statement that of all *n*-gons with the same perimeter, the regular one has the largest area, and so on.

For the teacher, the following historical and educational notes may be helpful:

The farming of the fertile field of isoperimetric problems has, of course, a long and interesting history (Porter, 1933). Euclid had already reserved a tiny portion of his *Elements* for talking about the largest areas enclosed by isoperimetric (equal perimeters) rectangles (Stowasser, 1976a). As Pappus of Alexandria (c. A.D. 320) said, Zenodor (c. 180 B.C.) proved some statements from the lost book about isoperimetric figures for comparison of polygons having equal perimeters with regular polygons and the circle (Pappus, 1932).

Various isoperimetric problems were solved much later—since 1694—by the brothers Jakob and Johann Bernoulli with analytic approaches (calculus of variation). Elementary geometric solutions were less asked for after Steiner. The so-called Fagnano problem, for example, done by analytic means had elegant elementary geometric solutions prepared first by H. A. Schwarz (1843-1921) and later by L. Fejer (1880-1959) (Courant & Robbins, 1941; Rademacher & Toeplitz, 1957; Stowasser, 1976b). A comprehensive compilation of the elementary geometric opti- mization approach was prepared by S. L'Huilier (1789). After Steiner's death, Polya was especially attracted to adapting Steiner's elementary method for his educational objectives (Polya, 1954, chap. 10). Also, various books based on Steiner's "World of Ideas" have been prepared so they can be understood by amateurs (Courant & Robbins, 1941; Dörrie, 1958; Hasse, 1955; Rademacher & Toeplitz, 1930). In the isoperimetric problem area, deep excavations will reward variation calculators and those who research convex solids. Newer developments are presented with relative simplicity by Yaglom and Boltjanski (1961), Lyusternik (1963), and Eggleston (1969). Blaschke (1956) is always worth reading. Through his isoperimetric proofs, Jakob Steiner has tacitly provided (unnoticed?) the existence of a maximal curve. Dirichlet indicated this gap in Steiner without much resonance. At age 70, Karl Weierstrass closed this existence gap in his lectures with the help of analytic methods. Without them, Caratheodory and Study first arrived at results in 1909.

Appendix

Historically Relevant Extreme Value Problems Solved Elegantly by Means of the Level-Line Method

R. J. K. STOWASSER

The two-dimensional isoperimetric problem was the subject of the preceding chapter. By specializing in isoperimetric rectangles, and with suitable alterations, elementary and elegantly solvable problems result, producing, moreover, an important general solution method (its analytical form is called "Lagrange's multipliers for extrema with constraints"). Of course, the ancient Greeks knew how to solve the following isoperimetric problems.

Problem 1A

GEOM

Among all isoperimetric rectangles, find the one with the greatest area.

Our "modern" solution uses the Cartesian coordinates system. We put the isoperimetric rectangles in the coordinates' corner and see that the upper-right vertices lie on a straight line with slope −1. ⑥

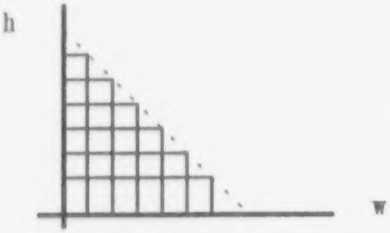
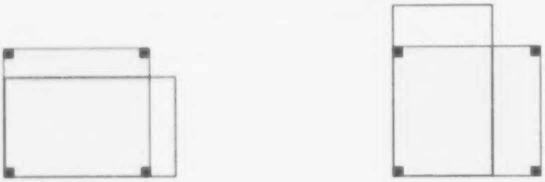


Figure 16.14

The increment of the area with growing $w < h$ is obvious and produces the rigorous proof that for $w = h$ the area of the rectangle becomes maximal.



The strip on the right has less area than the upper strip

The upper strip has less area than the strip on the right

Figure 16.15

The classical case is not very interesting. One can literally "see" the solution because of the symmetries. Altering problem 1A a little is much more fascinating.

Problem 2A

Along a straight river, cut a rectangular piece of land of maximal area and bounded by a given length of fence. (Along the river you do not need a fence, of course!) GEOM

The Cartesian representation looks like this:

⑥

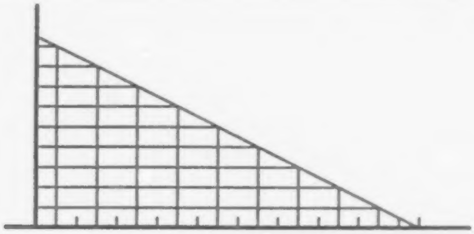


Figure 16.16

The "characteristic line" for the upper-right corner is a straight line with slope $-1/2$. Now you cannot "see" the maximal area rectangle any longer.

One has to do some calculations, appropriately with a calculator or a computer; for example, for $w + 2h = 360$ m:

⑫

w (meters)	...	170	...	178	180	182	190	...
h (meters)	...	95	...	91	90	89	85	...
A (m^2)	...	16,150	...	16,198	16,200	16,198	16,150	...

Consequently, the piece of land at the river should be twice as wide as it is long: $w = 2h$. The rigorous proof is analogous to the one in problem 1A.

We quickly solve another extreme value problem for the rectangles on the river.

Problem 3A

GEOM

Which of the "river rectangles" has extreme diagonal length?

Solution

We run along one of the perpendiculars to the characteristic line.

⑫

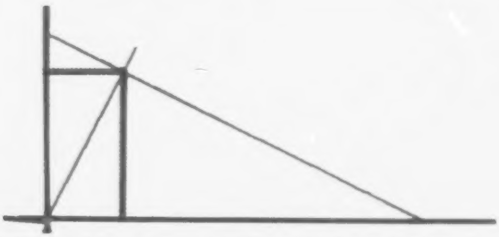


Figure 16.17

Simple similarity arguments prove: The "river rectangle" with minimal diagonal length runs away from the river twice as far as it is wide.

With problems 2A and 3A we can enter into the level-line method for geometrically solving extreme-value problems. For the "river rectangles" with $w + 2h = \text{constant}$ we have as level-line field the segments between the positive x-axis and the positive y-axis with slope $-1/2$.

For rectangles with constant diagonal length, the level-line field consists of quadrants of circles around the origin of the coordinate system which lie in the first quadrant. (For further extreme-value problems, one should also consider rectangles equal in area. Their level-line field consists of equilateral hyperbolas.)

The solution of problem 3A now takes the tangential circle to the characteristic line $w + 2h = 360$.

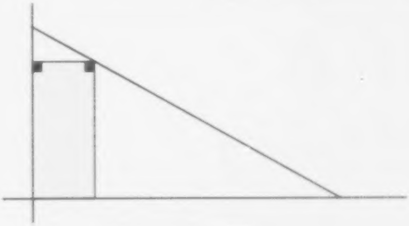


Figure 16.18

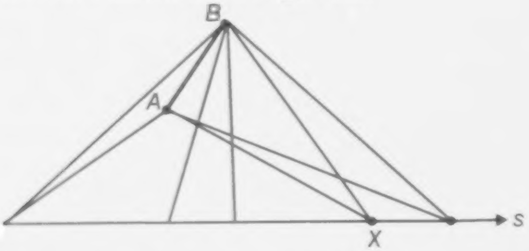
The level-line method, shown here in an extremely simple case, is very useful in many other extreme-value problems.

Problem 4A

A problem of Regiomontanus: Given a segment *AB* and a straight line *s*. Find the point *X* on *s* with maximal angle *AXB*.

GEOM

⑥



(*Hint:* Find the loci for the points *Y*, for which the measure of angle *AYB* is constant.)

Figure 16.19

Problem 5A

GEOM

A problem of Galileo Galilei: Given a curve c and a point P. Find the "inclined plank" on which a ball at rest at P will reach the curve in the shortest time.

⑥

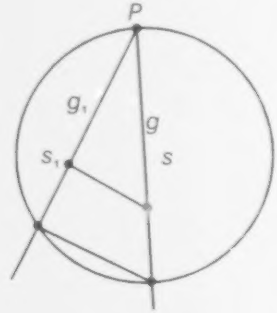


Figure 16.20

Solution

We draw a circle with P at its "highest" point and enter the acceleration g of the falling body and the acceleration g_1 operating along the slope, which is the perpendicular projection of g. S or s_1 mark the distances of the free fall or along the slope, respectively.

Because of $s = 0.5gt^2$ and $s_1 = 0.5\,g_1\,t^2$ and because of the right angle in the semicircle, $s_1/s = g_1/g$ is true.

This means that the level lines for the constant time are circles with P as their "highest" point, and therefore the following drawing shows the solution of the extreme-value problem 5A.

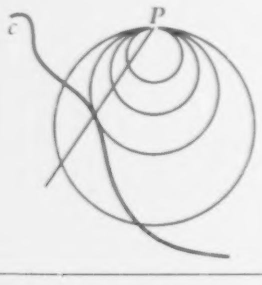
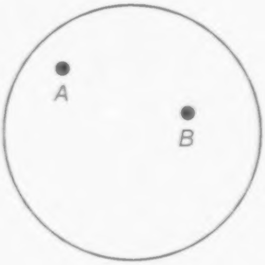


Figure 16.21

Problem 6A

The billiards problem of the Arab Alhazen: Billiards is played on a circular table. The ball at *A* should be hit so that after reflections off the table it hits the other ball at *B*.

GEOM



⑥

Figure 16.22

Solution

Here we use the Heron-Fermat principle, according to which the course a billiard ball takes from *A* to *B* with reflections on its way has an extreme length. We use the level-line field, which describes constant lengths of a stretch with a "bend." An ellipse defines equally long stretches with a bend starting from the foci *A* and *B*.

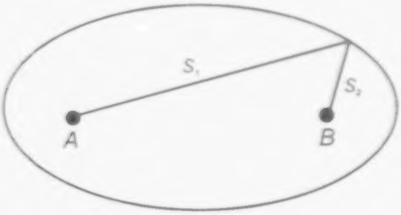


Figure 16.23

Therefore, in this case, the confocal ellipses with *A* and *B* as foci form the level-line field. The ellipses touching the circle (the round table) produce the solutions for the Alhazen problem.

Remark

In case *A* and *B* are equidistant from the center of the circle, there is a very simple solution without using the Heron-Fermat principle:

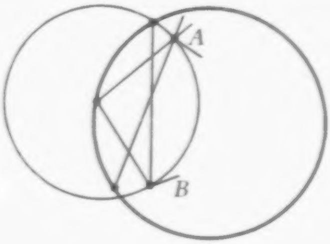


Figure 16.24

Figure 16.24 shows the three possible solutions. The proof is given by the circumferential angle theorem. (Think about it yourself!)

References

Blauschke, W. (1956). *Kreis und Kugel.* Berlin.

Courant, R., & Robbins, H. (1941). *What is mathematics?* New York: Oxford University Press.

Dörrie. (1958). *Triumpf der Mathematik.* Würzburg: Physica Verlag.

Eggleston, H. G. (1969). *Convexity.* Cambridge, UK: Cambridge University Press.

Hasse, H. (1955). *Proben mathematischer Forschung in allgemeinverständlicher Behandlung.* Frankfurt.

L'Huilier, S. (1789). *Polygonométrie et Abrégé d'Isopérimétrie élémentaire.* Geneva.

L'Huilier, S. (1782). *De relatione mutua capacitatis ex terminorum figurarum.* Warsaw.

Lyusternik, L. A. (1963). *Convex figures and polyhedra.* New York.

Pappus d'Alexandrie. (1932). *La Collection mathématique* (Ed. Paul Ver Eecke). Paris.

Polya, G. (1954). *Induction and analogy.* (Vol. 1 of *Mathematics and Plausible Reasoning*). Princeton, NJ: Princeton University Press.

Porter, T. J. (1933). *A history of the classical isoperimetric problem: Contributions to the calculus of variations.* Chicago: University of Chicago Press.

Rademacher, H., & Toeplitz, O. (1957). *The enjoyment of mathematics.* Princeton, NJ: Princeton University Press.

Steiner, J. (1971). *Gesammelte Werke* (2nd ed., Ed. K. Weierstrass, 2 vols.). Bronx, NY: Chelsea.

Stowasser, R. J. K. (1976a). Extremale Rechtecke: Eine Problemsequenz mit Kurzfilmen. *Der Mathematikunterricht,* 3.

Stowasser, R. J. K. (1976b). Küstenschiffahrt, Landmessen, Billard-Drei Problemfelder der Geometrie. *Der Mathematikunterricht,* 3.

Yaglom, J. M., & Boltjanski, W. G. (1961). *Convex figures.* New York: Holt, Rinehart, and Winston.

17

The Problem of the Duplication of a Cube

HANS K. KAISER

Problem solving is an art. This art has a lot to do with skills such as creativity of mind, flexibility of thinking, experience in the application of mathematical techniques, and the like. How can we learn this art? One way of mastering this art is "learning by doing," but much insight can also be gained by analyzing the way particular problems have been solved by famous mathematicians throughout the history of mathematics. From this point of view, one of the most instructive examples is the history of the famous problem of the duplication of a cube. In the following, we sketch a few important episodes from the history of this problem.

The problem of the duplication of a cube is the following:

> Given a cube. Construct one side of a cube which has a volume twice as large as the given cube.

GEOM

This problem is one of the three classical problems of antiquity (the other two problems are the trisection of any angle and the quadrature of the circle).

Eutocios, one of the later commentators of Archimedes, tells us something about the origin of the duplication problem by citing a letter from Eratosthenes to King Ptolemy. It starts as follows:

> It is said that one of the ancient tragic poets brought Minos on the scene, who had a tomb build for Glaucus. When he heard that the tomb was a hundred feet long in every direction, he said:
>
> "You have made the royal residence too small, it should be twice as great. Quickly double each side of the tomb, without spoiling the beautiful shape."
>
> He seems to have made a mistake. For when the sides are doubled, the area is enlarged fourfold and the volume eightfold. The geometers then started to investigate how to double a given body, without changing its shape, and this problem was called the duplication of the cube, since they started with a cube and tried to double it. After they had looked for a solution in vain for a long time, Hippocrates of Chios observed that, if only one could find two mean proportionals between two line segments, of which the larger one is double the smaller, then the cube would be duplicated. This transformed the difficulty into another one, not less great.
>
> It is further reported that, after some time, certain Delians, whom an oracle had given the task of doubling an altar, met the same difficulty. They sent emissaries to the geometers in Plato's academy to ask them for a solution. These took hold with great diligence of the problem of constructing two mean proportionals between two given lines. It is said that Archytas solved it with half cylinders, Eudoxus with so-called curved lines. (van der Waerden, 1963, p. 160)

④ It is most likely that this letter is a fake. But it tells us something about problem solving. If we cannot solve a given problem we may try to transform it into an equivalent problem and hope to be able to give a solution of the reformulated problem. Let us take a look at the contribution of Hippocrates. He claims the following:

ALG Let a and b be two line segments. If it is possible for us to solve the continued proportion $a : x = x : y = y : b$ (e.g., if we are able to construct x and y such that the continued proportion holds), we can solve the duplication problem. Indeed, if we solve the above proportion, we obtain $x^2 = ay$ and $y^2 = bx$. Hence, we have $x^4 = a^2 y^2 = a^2bx$, which yields $x^3 = a^2b$. Now we only have to choose for a the side of the original cube and then we set $b = 2a$. Then $x^3 = 2a^3$, which solves our problem.

How would we solve the continued proportion today? The proportion yields two equations: $x^2 = ay$ and $xy = ab$. We solve this system of two equations geometrically: We interpret $x^2 = ay$ as a parabola and $xy = ab$ as a hyperbola. The solutions for x and y we are looking for are the coordinates of the point of intersection of the two curves (see Figure 17.1).

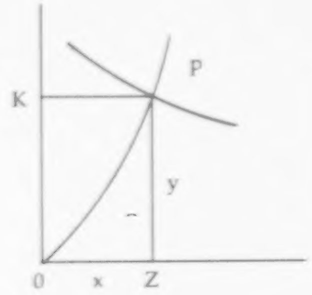


Figure 17.1

According to Eutocius's work, it was precisely this problem of the duplication of a cube that led to the invention of conic sections.

But, of course, there are several other possibilities of solving our continued proportion.

In the book of Eutocios we find the following solution, which is attributed to Plato (many historians of mathematics doubt that this attribution is correct). We start with an investigation of the figure below:

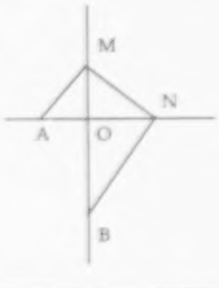


Figure 17.2

In a coordinate system we fix two points A and B as in Figure 17.2. Then we try to find two points M (on the y-axis) and N (on the x-axis) such that $\overline{AM}$ and $\overline{BN}$ are perpendicular to $\overline{MN}$. We claim that OM and ON are two mean proportionals between and OA and OB. Indeed, we observe that by the similarity of the triangles ΔOMA and ΔONM we get $OA : OM = OM : ON$. Again, by the similarity of the triangles ΔONM and ΔOBN we obtain $OM : ON = ON : OB$. Hence, we have $OA : OM = OM : ON = ON : OB$, which is what we wanted to show.

In order to construct *M* and *N*, Plato (according to Eutocios) devised the following instrument:

GEOM

Let *FGH* be a carpenter's square with a moveable ruler $\overline{KL}$, which is always kept perpendicular to $\overline{FG}$ (see Figure 17.3).

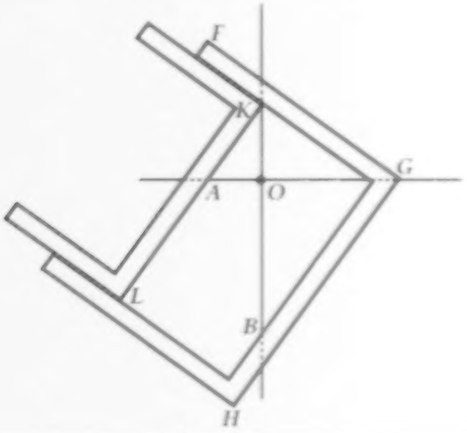


Figure 17.3

We draw a coordinate system and fix the point *A* on the *x*-axis and *B* on the *y*-axis. Then we position our instrument on the plane and move the ruler in such a way that *A* meets $\overline{KL}$ and *B* meets $\overline{GH}$, with *G* remaining on the line *AO* and *K* on the line *OB*. This solves our problem.

Another solution was proposed by Eratosthenes, a younger contemporary of Archimedes. He was so proud of this solution that he built a bronze model of his duplication instrument, which he put on top of a stone. And he gave orders to engrave the instruction for using it in the stone.

The instrument of Eratosthenes consists of three congruent rectangles that can be moved in a frame. On this frame we find a rod that can be rotated around point *O*. On each of the rectangles we draw the diagonals as shown in Figure 17.4.

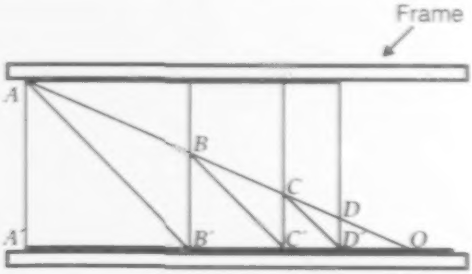


Figure 17.4

We reproduce the instruction for the use of the instrument as given by Eratosthenes:

> To determine two mean proportionals in continued proportion to two given lines. Let the lines AA' and DD' be given. Then I move the plates of the instrument towards each other until the points A, B, C, D lie on one line. Now, since AA' and BB' are parallel, OA and OB have the same ratio as OA' and OB'. And since AB' and BC' are also parallel, this ratio also equals that of OB' to OC'. Hence OA' is to OB' as OB' is to OC'. But AA' and BB' also have the same ratio, as well as BB' and CC'. In the same way we prove that BB' is to CC' as CC' is to DD'. Therefore, AA', BB', CC' and DD' form a continued proportion. Then two mean proportionals between two given lines have been found.
>
> If, now, the given line segments are not equal to AA' and DD' we shall obtain the mean proportionals by making AA' and DD' proportional to them; we reduce them to these and thus the task will have been carried out.
>
> In case more mean proportionals have to be found, we take every time one more plate in the instrument than the number of mean proportionals to be constructed. The proof remains the same. (van der Waerden, 1963, p. 231)

The mathematicians of ancient Greece did not accept the obtained solutions of the duplication problem as being mathematically correct. They claimed that the acceptance of such a solution would "destroy the beauty in mathematics." To understand this attitude we have to make an excursion into the history of mathematics in antiquity.

The birth of mathematics as a science took place in ancient Greece around 600 BC. The Greeks used mathematical methods, as did the Egyptians and the Babylonians before them. But while in Egypt and in Mesopotamia mathematics was performed by using (sometimes incorrect) recipes without explanations (as far as we know today), the Greeks tried to find arguments for the correctness of the various mathematical methods. In other words, ancient Greece introduced the concept of proof into mathematics. The first proofs of mathematical assertions are attributed to Thales of Milete. Unfortunately, we do not know how he proved mathematical results; we only know what he proved. Proclus reports that Thales was the first to prove (a) that a circle is divided into two equal parts by its diameter and (b) that when two straight lines intersect, angles are equal.

In the generation after Thales, the most eminent mathematician was Pythagoras of Samos. He used mathematics as his philosophical credo and tried to explain the world by mathematics. The doctrine of Pythagoras and his followers—the so-called Pythagoreans—proclaimed that God had ordered the universe by means of numbers. Harmony is divine and consists of numerical ratios. If someone achieves full understanding of this number harmony, he becomes divine and immortal.

The Pythagorean belief was confirmed by the theory of music. The Pythagoreans discovered that the most important consonant intervals in music could be obtained by shortening a string according to ratios of the numbers 1, 2, 3, and 4.

In modern terms, we can describe the number system used by the Pythagoreans as the system of positive integers to which the theory of proportions was applied. So their number concept was equivalent to our use of positive rational numbers.

Now, the diagonal of the unit square has the length $\sqrt{2}$, which is an irrational number, and this fact was already known to the early Pythagoreans. To be more precise, they could not explain the obviously existing object "diagonal of a unit square" by using their mathematical tools. In a way, this led to a first crisis on the foundations of mathematics. One way out of this dilemma is to apply the tool of proportions infinitely many times or, in modern terms, to approximate $\sqrt{2}$ by rational numbers. But this meant assuming that space can be divided infinitely many times. This was ridiculed by the philosophers like Zenon of Elea by creating paradoxes. One of these paradoxes is the paradox of the arrow: If we want to shoot an arrow from point A to a different point B, the arrow has to reach the midpoint M between A and B before it reaches B. In order to reach M it has to reach the midpoint M_1 of $\overline{AM}$, and so on. If we assume that space is infinitely divisible, the arrow can never move. This obviously contradicts our experience. Hence, the mathematicians do just nonsense which may contradict itself. So there had to be found another solution. On one hand, the Greek mathematicians started to consider arithmetical procedures in a geometric form, thus creating geometrical algebra. If one wanted to add two magnitudes, they had to be of the same kind (both areas or lines or solids). For example, addition of two lines was juxtaposition of the two lines, or the distributive law $a(b + c + \ldots) = ab + ac + \ldots$ was formulated by Euclid, in his *Elements*, in the following way:

> If there be two straight lines, and one of them be cut into any number of segments whatever, the rectangle contained by the two straight lines is equal to the rectangle contained by the uncut straight line and each of the segments. (Bk. II, Prop. 1)

On the other hand, Plato made the following proposal for making mathematics immune to criticism in the fashion of Zenon. He suggested fixing a number of postulates that are generally accepted to be true. From these basic postulates other propositions should be derived by using logical principles.

We find this idea, which we call the axiomatic method today, magnificently manifested in Euclid's *Elements*. Euclid formulates the following postulates:

> Let the following be postulated:
>
> 1. To draw a straight line from any point to any point.
> 2. To produce a finite straight line continuously in a straight line.

3. To describe a circle with any center and distance.
4. That all right angles are equal to one another.
5. That, if a straight line falling on two straight lines makes the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angles are less than two right angles. (Heath, 1956, p. 154)

From these postulates Euclid built the mathematical propositions of what we now know as "Euclidean geometry" step by step. For example, proposition 47 of book I is the famous Pythagorean theorem.

We return to our duplication problem. Of course, we look for a solution that employs only construction steps on the basis of Euclid's postulates. The relevant postulates are numbers 1 through 3. We may translate the relevant postulates into the following picture: We are allowed to use an unmarked ruler and a compasses. To be more precise, looking at the postulates does not make immediately clear the solution to the following problem:

> Given two unequal straight lines, to cut off from the greater a straight line equal to the less.

But in Book I, Proposition 3, Euclid gives a proof that such a construction is possible. So we can say that a mathematically correct construction is one which makes use of only an (unmarked) ruler and a compass and which consists of only a finite number of construction steps. The use of other instruments is prohibited. So the solutions of both Plato and Eratosthenes cannot be accepted as correct, at least not in the way they have been presented here. And our solution via conic sections cannot be accepted either, since, with Euclidean tools, we can only construct discrete points of the curve.

Therefore, Greek mathematicians tried again and again to find a solution of the duplication problem by using Euclidean tools only. But they did not succeed. A good proportion of the higher geometry in antiquity was invented in these various unsuccessful attempts. As an example, we discuss the solution of the duplication problem which makes use of the so-called cissoid (introduced by Diocles):

Let $\overline{AB}$ and $\overline{CD}$ be two mutually perpendicular diameters of a circle. Let E and Z lie on either side of, and at equal distances to, B. Draw $\overline{ZH}$ perpendicular to $\overline{CD}$ and let $\overline{ED}$ meet $\overline{ZH}$ in X. As E and Z move along the circumference of the circle, X describes a curve, which is called cissoid (see Figure 17.5).

GEOM

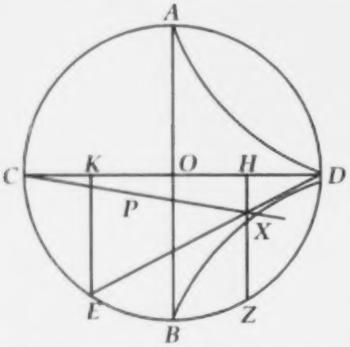


Figure 17.5

One may use this curve to find two mean proportionals between given line segments.

ALG

We draw $\overline{EK}$ perpendicular to $\overline{CD}$. The similarity of the triangles ΔDXH and ΔDEK yields $DH : HX = DK : KE$. Since $DK = CH$ and $KE = HZ$, we have $DH : HX = CH : HZ$.

Observe that ΔDZC is rectangular, hence we obtain $HZ^2 = CH \times DH$, or $HZ : DH = CH : HZ$. Putting these two proportions together gives $CH : HZ = HZ : DH = DH : HX$.

Thus, HZ and DH are two mean proportionals between CH and HX. In order to solve our duplication problem, we have only to find X on a given cissoid such that $CH : HX = 2 : 1$.

This can be achieved by locating the intersection of the cissoid on the line CP that determines on the two mutually perpendicular radii OC and OB segments OC and OP in the given ratio 2 : 1 (again, we make use of the similarity of triangles: ΔOPC is similar to ΔHXC). Again, this solution was unacceptable from Euclid's point of view, since we only know a pointwise construction of the cissoid under the sole use of Euclidean tools.

Today we know that the Greek geometers were bound to fail in their efforts to solve the duplication problem on the basis of Euclid's postulates. It is not possible to find a construction for doubling a cube by using Euclidean tools only. We are going to outline the proof of this proposition. We regard the problem from a different point of view; namely, we shall look at the geometrical problem from the algebraic point of view and employ a little bit of modern algebra.

ALG

Suppose we are given a line segment which we define to be one unit length. A real number α is called constructible if we are able to construct a line segment $|\alpha|$ in a finite number of steps from this unit length by using only ruler and compasses.

If we are given constructible numbers α and β, then $\alpha + \beta$ is constructible as well. For $\alpha, \beta > 0$ the proof is given geometrically:

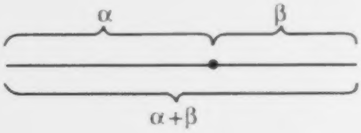


Figure 17.6

The other possible cases, $\alpha > 0, \beta < 0$; $\alpha < 0, \beta > 0$; $\alpha < 0, \beta < 0$, are shown in a similar way.

Similarly, we proceed if we want to show that $\alpha - \beta$ is also constructible. The next two figures indicate the following result: If α and β are constructible, then $\alpha \times \beta$ and $\alpha \times 1/\beta$ $(\beta \neq 0)$ are constructible.

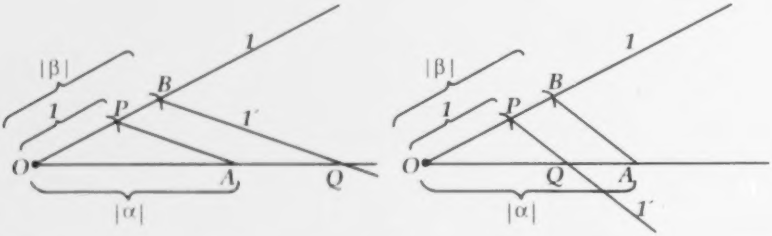


Figure 17.7

A modern way of formulating the combined result is to say: The set of constructible numbers forms a subfield S of the field R of real numbers.

Now we know that the smallest subfield contained in R is the field **Q** of rational numbers. Hence we have $S \geq \mathbf{Q}$; in other words, every rational number is constructible. Therefore, we can construct any point (q, p) in the plane with $q, p \in \mathbf{Q}$. Any further point in the plane which can be located by using compasses and ruler can be found in the following three ways:

1. as an intersection of two lines, each of which passes through two known points having rational coordinates,
2. as an intersection of a line which passes through two points having rational coordinates and a circle whose center has rational coordinates and the square of whose radius is rational,
3. as an intersection of two circles whose centers have rational coordinates and the square of whose radii are rational.

Let us take a look at these three construction possibilities algebraically.

The first is just the task of finding the solution of the system of equations that represent the two lines

$$\begin{aligned} &ax+by+c=0 && (a,b,c\in \mathbf{Q})\\ &a_1x+b_1y+c_1=0 && (a_1,b_1,c_1\in \mathbf{Q}). \end{aligned}$$

The second may be viewed as finding the solutions of the system:

$$\begin{aligned} &x^2+y^2+dx+ey+f=0 && (d,e,f\in \mathbf{Q})\\ &ax+by+c=0 && (a,b,c\in \mathbf{Q}). \end{aligned}$$

The third is of the same type as the second, since it may be considered as the intersection of one circle with the common chord of the two circles, algebraically: the solution of the system consisting of the equations of the two circles

$$\begin{aligned} &x^2+y^2+d_1x+e_1y+f_1=0 && (d_1,e_1,f_1\in \mathbf{Q})\\ &x^2+y^2+d_2x+e_2y+f_2=0 && (d_2,e_2,f_2\in \mathbf{Q}) \end{aligned}$$

is the same as

$$\begin{aligned} &x^2+y^2+d_1x+e_1y+f_1=0\\ &(d_1-d_2)x+(e_1+e_2)y+(f_1+f_2)=0. \end{aligned}$$

Since **Q** is a field, the solution of case 1 yields just a pair of elements of **Q**. Case 2 leads essentially to the solution of a quadratic equation with coefficients in **Q**. Such an equation, when solved by the quadratic formula, may have solutions involving square roots of numbers which are not squares in **Q**. If this is the case, we obtain—since we know that the constructible numbers form a field—the smallest field which contains **Q** and $\sqrt{\alpha}$. We denote this field by $\mathbf{Q}(\sqrt{2})$.

If we have already a field F of constructible numbers, and we obtain by step 2 a "next new number," then this number lies in the field $F(\sqrt{\alpha})$ for same $\alpha \in F$, $\alpha > 0$.

The following figure shows that the square root $\sqrt{\alpha}$ of a constructible number $\alpha > 0$ is again constructible ($\sqrt{\alpha} = OQ$):

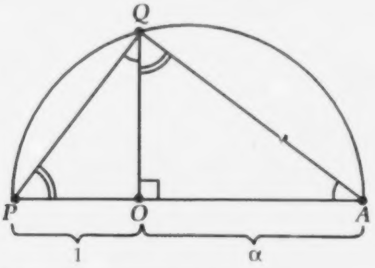


Figure 17.8

Hence, we have shown the following important result: The field F of constructible numbers consists precisely of all real numbers which can be obtained from **Q** by taking square roots of positive numbers a finite number of times and applying a finite number of field operations.

In modern algebraic terminology, we can formalize this result as follows: If α is constructible and $\alpha \notin \mathbf{Q}$, then there exists a finite sequence of real numbers $\alpha_o \in \mathbf{Q}, \alpha_1, \ldots, \alpha_n = \alpha$ such that $\mathbf{Q}(\alpha_1, \ldots \alpha_i)$ is an extension of $\mathbf{Q}(\alpha_1, \ldots, \alpha_{i-1})$ of degree 2, $i = 1, 2, \ldots, n$. In particular, the degree of $\mathbf{Q}(\alpha)$ over $\mathbf{Q}$: $[\mathbf{Q}(\alpha) : \mathbf{Q}] = 2^r$ for some integer $r \geq 0$. (Degree 2 means that α_i is the zero of a polynomial of degree 2, which is irreducible over $\mathbf{Q}(\alpha_1, \ldots, \alpha_{i-1})$.)

With this machinery at hand, we can easily show that the problem of the duplication of a cube cannot be solved with ruler and compasses alone: Let the given cube have a side of length 1. Hence it has volume 1. The cube we look for has volume 2, hence a side of length $\sqrt[3]{2}$. But $\sqrt[3]{2}$ is a zero of the irreducible polynomial $x^3 - 2$ over **Q**. Hence the degree of $\mathbf{Q}(\sqrt[3]{2})$ over **Q** is 3. In view of the above result, we would need to have $3 = 2^r$ for some integer r. Clearly, such an integer r does not exist.

Our perspective on problem solving is certainly enriched by our little excursion into the history of the duplication of the cube problem. Most people consider mathematics as a natural tool for solving problems. But there is not always an immediate solution at hand, and it is not always immediately obvious which method we should apply. One way of coping with such a situation is to look at the problem from different viewpoints or, to put it in a more mathematical language, try to transform the problem into an equivalent one. We then try to solve the problem in its new form. We find the approach not only in the history of the duplication of the cube problem but also, for example, in modern times in various attempts to solve the famous four-color map problem. The mechanical solutions of our problem are nice examples of recursive thinking, which is another technique often used in problem solving. One of the main features of the story told above is the question of accepting a solution as correct or not. In our case, the attempts at finding solutions with Euclidean tools only brought forth most of the higher geometry of antiquity. The proof of the impossibility of a solution on the basis of Euclid's axioms again is an example of changing from one point of view (geometrical modeling of the problem) to another one (algebraic modeling of the problem). So, we see, problem solving asks for a number of skills, of which flexibility of thinking and creativity are not the least.

Reference

van der Waerden, B. L. (1963). *Science awakening*. New York: John Wiley.

Suggested Reading

Dorninger, D., & Müller, W. B. (1984). *Allgemeine Algebra and Anwendungen.* Stuttgart: B. G. Teubner.

Eves, H. (1983). *An introduction to the history of mathematics.* New York: Saunders College Publishing.

Gilbert, W. J. (1976). *Modern algebra with applications.* New York: John Wiley.

Heath, T. L. (1956). *Euclid: The thirteen books of the elements.* New York: Dover.

Kaiser, H. K., & Nöbauer, W. (1984). *Geschichte der Mathematik für den Schulunterricht.* Vienna: Hölder, Pichler, Tempsky.

18

Solving Mathematical Problems Using Game Strategies

MARION KAUKE

SABINE ZILLER

Introduction: Task Solving by Algorithms

It is widespread practice in the teaching of mathematics to pose problems, teach a solution scheme, and then examine students on this particular knowledge. This method may appear to be economical, but it cannot be sufficient to prepare children adequately for the demands of life. Children hardly learn to discover problems by themselves, to draft various paths to the solution, or to transfer patterns to similar, basically analogous tasks. Their ways of thinking are usually limited to the procedures seen in their experiences, which are often memorized.

It is surely right to demonstrate paths to solutions and impart knowledge to students, but a more promising approach is to enable students to learn independently through problem solving and developing their individual talents. One should not, of course, expose the children to an endless and unproductive search for solutions, but should introduce rational methods of thinking which build students' self-confidence.

This chapter will deal with the topic of how even younger children, aged 7 to 12, can be optimally motivated by successful classroom experiences to enjoy mathematical tasks, promote their skills by puzzling out solutions, and cultivate a permanent interest in mathematics. This kind of teaching will challenge, and perhaps inspire, *all* students, even those who are undisciplined, disheartened, or not so good. It is not aimed chiefly at the promotion of special talents.

Creating a Stimulating Atmosphere

Teachers should place high value on the indirect influences that help develop students' confidence in their ability to find a solution. The creation of a stimulating atmosphere for learning is still often neglected. The teacher sometimes begins immediately with abstract subject matter and urges students to proceed with the set task. No wonder the majority of children, who are distracted from their joy of playing by everyday school life, hardly want to follow the teacher's explanations.

Moreover, mathematics has such a formidable reputation that it can intimidate overmotivated, anxious children (symptoms of "math phobia" are described in specialized literature). It is necessary to alleviate this fear so students at least try to solve the tasks presented and do not lose the thread too fast or get bored because the abstract matter taxes their abilities too heavily.

Therefore, the first principle for a mathematics teacher who wants to motivate is to create a relaxing atmosphere.

First Principle: Create a Relaxed Environment

To "create a relaxed environment" means nothing more than to help students forget their anxieties, their usual stress in school and the stress to perform (time pressures, grades), and the fear of competition. Creating a relaxed environment also entails leading students into a world of imagination—the realm of numbers or figures. Whether the teacher uses an amusing anecdote or an aesthetically perfect picture, the aim is achieved if the teacher attracts the attention of the children so that they can eagerly anticipate a mental adventure.

This fascination presupposes some well-planned steps by the teacher. First, leisure is required, a mood that lets students know there is more than enough time to develop the tasks in depth, without fear of interruption or disturbance. "Relax and enjoy" is the most important message the teacher should radiate.

Furthermore, he or she should be an example of courage to the children to face a perhaps uncertain venture. Nobody will expect perfect solutions in the beginning. If the children are intimidated, one should praise a correct statement, even if it is very basic, as promising. Unusual approaches to the solution should not be dismissed but rather praised as particularly original. It is absolutely wrong to compare students' performance with the knowledge of the teacher or of the best student, but this does not preclude increasing step by step one's expectations for the mathematical abilities of the children.

A teacher with a good sense of humor has the talent to recognize when the trickiness of a problem has reached its limit and immediately interjects a humorous comment to defuse the students' frustration. The mental energy gained by this can be channeled into finding a successful solution. A clever teacher approaches the situation from the start by assuming the least common knowledge. In this respect, it makes more instructional sense to dampen motivation than to say, "Try a bit harder," since too much strain can cramp the cognitive aspect of learning. With a reliance on self-organization, one should not interfere with the natural process; rather, let it develop by itself.

So the pattern is changed from fear of failure to hope for success. The relative disassociation from immediate responsibility encourages the type of alienation that allows the problem to be seen in a different light, thus releasing students from preconceptions that can hinder creative problem solving.

This clearly is not a mystical luxury but a *conditio sine qua non* for improving students' "brain power" and mental energy. The capacity of healthy children increases, thus lessening boredom and sleepiness and ensuring students' optimal concentration on the task.

Second Principle: Using Imagination

After students feel pressures diminish, the teacher should bring them into the realm of mathematical imagination. This starts with students convincing themselves that although they may not be mathematics fanatics, they can *act* as if they were.

There are a lot of possibilities for stimulating such an attitude, mostly having to do with the symbol-substance of numbers, their vivid representation of numerical facts. The teacher could encourage students to use objects to represent numbers, such as baseball cards, jet planes, jewels, coins, video games, and so on. The idea of using Morse code also rouses motivational energies. The teacher can discuss how one can encode a language so that it is accessible only to "insiders." These methods of personalizing subject matter and boosting motivation can make students more enthusiastic and give classwork more personal importance.

It is always necessary to find out what occupies the students' minds in each age group. Armed with this knowledge, the teacher can find various motives and suitable wordings for the tasks, which would otherwise appear quite dry. Students should by no means be hindered if they want to put themselves in the position of an unbeatable detective or a commander whose power depends on the solution of exciting navigation problems. Students could also put a stop to a criminal's activities by cracking difficult codes and presenting evidence in a mock court. Mathematical puzzles during a treasure hunt (perhaps set in an imaginary world) can provide exciting adventures for children of different age groups. Humanitarian relief actions with mathematical calculations can inspire the altruistic engagement of the students. And using e-mail to solve problems with children in other countries can be an intriguing situation for students.

Psychologists have found that putting problems in context enables students—even the less talented—to solve geometric tasks that usually prove too difficult when presented as just abstract patterns. The teacher's imagination can put students in a mindset of building the Egyptian pyramids as the origin of civilization or of animating a "virtual" future world of computers. A dramatic biography of a well-known mathematician can promote affection for the subject. A teacher could use the history of mathematics or a related field to demonstrate how similar problems have been solved, before students try to "reinvent the wheel."

Mathematical fantasy does not have to be based in a more or less real past, present, or future; it could, for example, take the form of imagining that, with the help of a code word, students could call a mathematical genius who tells them the solution. At first, these games may seem nonsensical, but psychologically they are very effective. They can reduce the seriousness of the problem and relegate it to a higher power, an unconscious intuition, that sometimes becomes blocked or repressed.

Visualization is a method used by many top mathematicians to simplify the search for the thought that breaks through the barrier of a challenge. Something like this imagination now becomes accessible to many people, to some extent, through computers. For younger students, computer-generated games that translate abstract mathematical matters into vivid fantasies would be recommended.

The prospects of success can also be unlocked with the magic question, How would fairy-tale figures—a dwarf, a giant, or a princess—solve these tasks? Embedding problems in contexts remote from reality can bridge the difficulty of given problems and demanded solutions by translating problems into psychologically simpler terms.

Third Principle: Experimental Exploration of Solution Principles

③ It must be granted that all students can make themselves *spontaneously* familiar with the tasks, their circumstances, and their demands. Everybody should find his or her own pace and degree of difficulty. Instead of reverential understanding of the task, students should confidently deal with the parameters of the problem in an experimental way, which means nothing more than playfully initiating, varying, repeating, and controlling conditions.

This procedure will at first be based on trial and error, but it will lead to success in understanding. The "tricks" of asking the right questions will be recognized through playful trying. Emphasized and combined, elements of game strategies soon form a pattern: the model of the adequate solution.

Such solution patterns are to be found in various ways in the demands of games. It is well known that a "double row" in the traditional game of ancient Roman times is of considerable strategic advantage in shortening the way to victory.

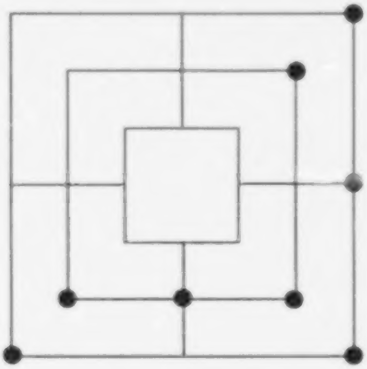


Figure 18.1

In a broad sense, this conceals a principle of creativity called *multiple usage,* which means a strategy can remain viable even under certain contingencies. In the game of chess, combinations are typical phenomena of creative thinking. A desirable solution demands precise coordination of the matching figures and circumstances. In many cases, the player has to sacrifice an important figure, a figure one would normally try to keep by all means. But the risk of losing the piece is not real. A well-calculated sacrifice is a move that gives up material to gain a tactical advantage. The principle of combination as a hidden instrument of efficiency can be found in many areas of life.

It has been known for some time that long-range thinking, economy of thoughts, and soundness and efficiency of trains of thought can be effectively improved by systematic chess games (Bönsch, 1987), as can characteristics of creative thinking, such as originality, fluidity, and flexibility (Kauke, 1992). Chess, the king of brain twisters, contains a variety of tasks and potential problems. They are a special reflection of, and at the same time a good possibility for, preparation for the challenges of life. (Klaus, 1969, has noted that in light of the information explosion, teaching needs to change its focus from the rote learning of facts to the learning of methods for obtaining and interpreting facts.) It is troublesome, if not impossible, to calculate the number of promising or misleading moves when playing chess. Because of this, the player tries to simplify the mental search or cut the selection process. A time-honored rule—for example, when a player ponders risking her advantage by leaving her opponent's piece on the board—is therefore "When in doubt, take it!" In this way, one removes the possible threat. This procedure corresponds to the principle of creative "cutting" and is also adapted for heuristic search strategies in programming chess-playing computers. Posting the figures on efficient fields, where they simultaneously cover each other and attack the opponent's forces, doubles the efficiency of playing. We recognize it as a case of the fundamental mathematical procedure of addition. In a broader sense, it can be seen as a multiple usage of possibilities in order to increase the chances of overcoming a variety of problems. The effect is even greater when a player succeeds in changing a pawn in the last row into a queen. In general, a pawn possesses the value of one unit, and the queen, nine units, so that a player realizes a ninefold gain in efficiency, which we can compare with the mathematical principle of multiplication. Conversely, the operation of division can be used to disperse the opposing constellations of figures and gain an advantage. On the other hand, a clumsy division of one's own forces can decrease one's own potential. These are only the first steps in creating efficiency in solving mathematical problems.

There are a great number of intellectually rigorous games that demonstrate the general principles of creativity. For example, jigsaw puzzles require students to create a pictorial whole by coherently joining bits and pieces. There are plenty of children who can play cards ingeniously but who apparently have almost insurmountable problems when asked to solve analogous algebra problems. Once recognized, there are many patterns that occur, sometimes in modified form, in mathematical tasks and puzzles. If one already knows some key principles from game practice, the solution patterns can be applied to analogous tasks and in various similar fields of application. Those seemingly frivolous ways of motivating students are handsomely compensated for by the educational benefit.

Instruction by Creative Heuristics

Creative heuristics are a way of learning through trial and error so as to shorten the time it takes to get to the solution. More judicious teaching strategies do not accept as a rule of thumb the notion that we learn best how to solve problems by solving them. Instead, they try to develop rational systems for finding ideas to solve the tasks; more exactly, these heuristic methods help a player in the search for creative ideas and direct a player out of the "labyrinth of uncertainties." One can borrow here not only from mathematical game theory but chiefly from the experience of players, engineers, and technicians. In fact, there are a few dozen common principles that form the basis of ideas for most modern inventions.

Relatively elaborate strategies for solving problems play an important part in this book. These strategies are to be found in the introduction (see also Posamentier & Stepelman, 1995, pp. 115-119). They can be introduced with examples from games or emotionally appealing texts, or the teacher can ask the students to apply the strategies used in such games as checkers and chess. The teacher can also try to emphasize the heuristic principles and exercise them in modified contexts. There are no limits to how a talented mathematics teacher can teach creative search principles. He or she should generally introduce mathematical tasks in such a way that the students come to value not only a successful solution but also the ability to articulate how they achieved it.

The mathematics teacher should also show students the advantage of expanding the range of means to the right solution. A narrow path to solving a problem can lead to psychological barriers, which hinder productive solutions. ③

An important way of eliminating psychological barriers is to tolerate even seemingly inadmissible things: Just saying "That's not possible" only testifies to rigid thinking and destroys the train of thought. It is more constructive to ask the student to follow through on that train of thought.

It is important to show how a problem can be transformed into equivalent terms, or without losing its essential features, until one comes to a familiar path to the solution.

Highly individual thinkers owe their originality and flexibility to free ways of thinking and to their courage in thinking. Teachers should vividly demonstrate this to students and allow them to jump from one idea to another in problem solving.

The following tasks illustrate creative heuristics as part of teaching strategies. Fairy tales in particular are adapted to children's mentality and are suitable for the introduction, abstraction, and exercise of creative principles.

These examples are such that every student can find at least one solution or one statement of the solution. The students' different approaches should then serve as the starting point for becoming aware of the intuitive procedure, the finding of strategies, and recognizing patterns. These help to systemize and facilitate the search for solutions. Everybody should get a chance to try the tasks and test ideas.

Example 1

> In the fairy tale "The Devil With the Three Golden Hairs," the hero is given a task by the king: In order to be allowed to marry the king's daughter, he must find the three golden hairs of the devil. Arriving in hell, he asks the devil's grandmother to draw out the three golden hairs for him, which she does with cunning. ALG
>
> Now imagine the grandmother is blind, and to make the task easier, the hero has to bring only one golden hair. Since the devil has altogether five hairs on his head, three of which are the gold ones, how many hairs does the blind grandmother have to pull out to be sure she got at least one golden hair? COMB

The first principle for understanding the real problem is to consider the extreme case: that the grandmother first pulls out all nongolden hairs. ⑤

⑪ The second principle is to examine systematically all possible outcomes.

⑥ The third principle is to find the solution in a graphic way, by creating a picture or table.

⑬ To generalize the solution, deductive and inductive reasoning are important. One can generalize the task from 1, 2, 3, . . . hairs to finally x hairs on his head, and
⑱ from 1, 2, 3, . . . golden hairs to finally y hairs. Then the task can be generalized this way: to pull out at least 1, 2, 3, . . . hairs to find a general formula for x hairs, of which at least y hairs have to be golden and z of these y golden hairs have to be pulled out. Find the necessary relations between x, y, and z.

Solution

Let us assume the most general case:

1. The devil has x hairs on his head,
2. y of them are golden.

The task then becomes: How many hairs does the blind grandmother have to pull out of the devil's head to make sure she got at least z golden hairs?

First of all, for the task to be valid, $y \le x$ and $z \le y$, and all values have to be positive integers, where upon $x, y, z > 0$.

The extreme case is that the blind grandmother first pulls out all nongolden hairs; that means she does not get one golden hair in the beginning.

From assumptions 1 and 2 it follows that the devil has $x - y$ nongolden hairs on his head. So if the grandmother pulls out $x - y$ hairs, the worst case is that these could be all nongolden. In this case, the $(x - y + 1)^{th}$ hair is definitely a golden one, and therefore so are all succeeding hairs.

To get at least z golden hairs, the grandmother has to pull out $x - y + z$ devil's hairs.

(*Note:* The task can also be made more difficult by leaving out the number of hairs on the devil's head. The students then have to find out themselves that this value is necessary in order to find a solution.)

A large variety of similar tasks can be formulated and solved.

Example 2

ALG
COMB

We are in the realm of natural numbers, and we now want to have a closer look at the 2-digit numbers in that realm. Outwardly regarded, each 2 digits, pairs of the digits from 0 through 9, constitute a pair. How many 2-digit numbers are there?

Let us suppose only the digits 0, 1, 2, and 3 are available. What 2-digit pairs can be built from them?

One might first solve simpler analogous problems by: ④

a. starting with only the digits 0 and 1 being admitted
b. determining how the solution changes if only the used/admitted digits play a part but their order is irrelevant
c. considering more graphic tasks, such as finding the number of possible ways there are to choose two fingers from one hand

The first principle for finding the solution is to account systematically for all possibilities. ⑪

Another principle is to find the solution graphically, by making a picture or table. ⑥

Again, deductive and inductive reasoning play a part in generalizing the solution. What does the problem look like for additional fours and so forth? What solution do we get if not pairs but groups of three are wanted? What further possibilities of generalization can be created? ⑬ ⑱

Solution 1

The number of all consecutive natural numbers between x and y, $x < y$, is $y - x + 1$.

Two-digit numbers are all between 10 (smallest 2-digit number) and 99 (largest 2-digit number). Since $99 - 10 + 1 = 90$, there are exactly 90 2-digit numbers.

Solution 2

The first digits can be 1, . . . , 9; the second can be 0, . . . , 9. Every possible first digit can be combined with every possible second digit; that means 1 as the first digit can be combined with 0, . . . , 9. ③

There are exactly 9 possibilities for the first digit and 10 possibilities for the second digit. Thus, there are altogether $9 \times 10 = 90$ combinations of digits to form a 2-digit number.

We could also make this a more general problem:

Find the quantity of all n-digit numbers ($n \geq 1$ be a natural number).

Solution 1

The quantity of all natural consecutive numbers between x and y, $x < y$, is $y - x + 1$. The n-digit numbers are all between 10^{n-1} (smallest n-digit number) and $9(10^0 + 10^1 + 10^2 + \ldots + 10^{n-1})$ (greatest n-digit number).

Since $9(10^0 + 10^1 + 10^2 + \ldots + 10^{n-1}) - 10^{n-1} + 1 = 10^n - 10^{n-1}(10 - 1) = 9 \times 10^{n-1}$, there are exactly $9 \times 10^{n-1}$ n-digit numbers.

Solution 2

③ The first figure can be 1, . . . , 9, the second, third, and so forth can be 0, . . . , 9.

Every possible first digit can be combined with every possible second digit and every possible third digit; that means 1 as the first digit can be combined with 0, . . . , 9 as the second digit (10 possibilities each) and with 0, . . . , 9 as the third digit. That makes already 9 × 10 × 10 possibilities.

There are exactly 9 possibilities for the first digit and 10 possibilities for each subsequent digit. There are exactly $n - 1$ subsequent digits. Thus, there are altogether $9 \times 10^{n-1}$ combinations to form an n-digit number.

There is a large variety of possibilities for formulating similar analogous tasks, such as using letters of the alphabet instead of numbers. The adoption of a different point of view—that is, finding the quantity of multidigit numbers by various ways—may increase the understanding of the solution.

Example 3

COMB

ARITH

What possibilities are there for disassembling this figure?

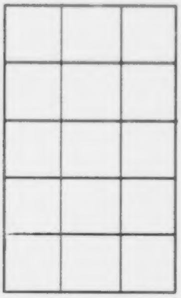


Figure 18.2

All forms of (connecting) box-figures are allowed.

What tasks (addition, subtraction, multiplication, division) can be expressed graphically by the taking apart and joining?

④ The first principle for finding the solution is, again, to solve simpler analogous problems.

The second principle is to solve it graphically, by making a picture or a table; one could, for instance, work with scissors and paper. ⑥

The third principle is to account systematically for all possibilities. ⑪

One could also find analogous tasks, both smaller and bigger, such as this: ④

Combine the following figures so as to obtain a rectangle:

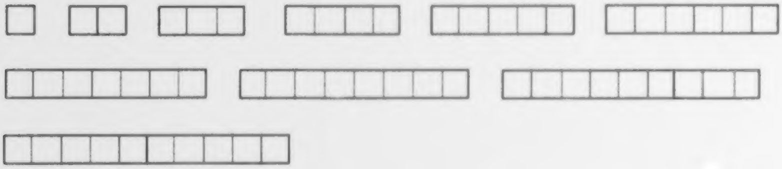
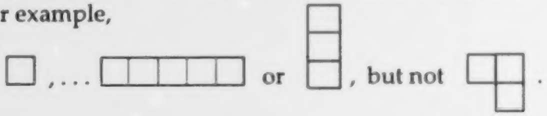


Figure 18.3

There is an amazing variety of possibilities for solving this problem. This holds true, for instance, for the graphical expression of addition/subtraction tasks, multiplication/division tasks, and linking up with tasks of division with remainders/congruences. There are almost no limits to generalization and to extensions. ⑱

As we cannot treat all variants, we introduce some constraints: First, only sequences of squares in a row or column contained in Figure 18.2 will be accepted as "pieces," for example,



Second, we will only consider separations into at most 5 such pieces.

Solution

Part 1: From the preconditions it follows that every piece must have a size between 1 and 5 squares (that means each summand must be between 1 and 5). It follows as well that there must be at least 3 pieces.

The only solution to separate the figure into 3 pieces is the one with 3 "fivers." For separation into 4 pieces there are the following possibilities: ⑪

5 + 5 + 4 + 1
5 + 5 + 3 + 2
5 + 4 + 4 + 2
4 + 4 + 4 + 3.

For the taking in five pieces:

5 + 5 + 3 + 1 + 1
5 + 5 + 2 + 2 + 1
5 + 4 + 4 + 1 + 1

5 + 4 + 3 + 2 + 1
5 + 4 + 2 + 2 + 2
5 + 3 + 3 + 2 + 2
4 + 4 + 4 + 2 + 1

4 + 4 + 3 + 2 + 2
4 + 4 + 3 + 3 + 1
4 + 3 + 3 + 3 + 2
3 + 3 + 3 + 3 + 3.

Part 2: We want to consider in particular the solution 5 + 4 + 3 + 2 + 1, as their summands are consecutive natural numbers beginning with 1. This allows us to draw parallels to the anecdote about the famous mathematician Gauss. He and his classmates were supposed to solve the task of finding the sum of 1 + 2 + . . . + 99 + 100. The teacher had probably hoped to get a quiet period by assigning this task. But he didn't, because Gauss found a quick and elegant solution for this task.

It is reasonable to pose the general task to join *n* pieces (rectangles), each formed by 1 through *n* units, or squares (depending on the age of the students, one could also use different colors), so that there is in the end again a rectangle.

The students can first find solutions themselves by trial and error: for example, all in one row, for 1, 2, 3, . . . , *n* rectangles.

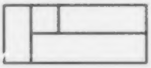
⑥ For 4 rectangles with 1 through 4 units:

Figure 18.4

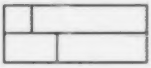
we get the following solutions:

a. all in a row

b.



c.



For 5 rectangles with 1 through 5 units:

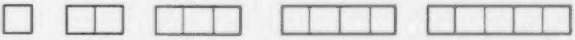


Figure 18.5

we get the same solutions as we got for 4 rectangles, but the strip of 5 is added.

For 6 rectangles with 1 through 6 units:

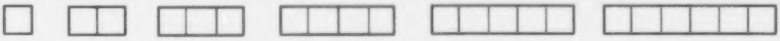
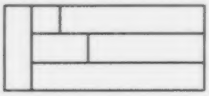


Figure 18.6

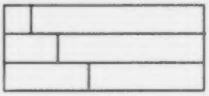
we get the following solutions:

a. all in a row

b.



c.



The following generalization can now be recognized: There is one solution for every *n* by joining the largest element with the smallest, and afterwards joining again the largest and the smallest element of the remaining pieces, and so on.

Grouping together and adding the respective external elements in this manner, we simply count the pairs to obtain as summing formula: ⑬

- for even n: $(n + 1)(n/2)$
- for odd n: with an element left over in the middle $\frac{(n+1)}{2}$
  $(n+1) \cdot \frac{(n+1)}{2} + \frac{(n+4)}{2} = \frac{n\,(n+1)}{2}$

Example 4

This is an example for the "wording" of a task in an (imaginary) fairy tale:

> The hero is sitting in the dungeon and is given the task of inventing a new flag, which does not already exist, for the ruler's country. If he does not offer at least *x* proposals, he shall lose his life. Can we help him?

COMB

The first principles for determining the scope of the solution are:

- organizing data ⑭
- approximating ⑮
- determining necessary and sufficient conditions ⑧
- determining characteristics of objects ⑯

A further principle for finding the solution is to account systematically for all possibilities. ⑪

Another principle is to find the solution graphically. ⑥

Again, deductive and inductive reasoning play a part in generalizing the solution. ⑱ ⑬

Solution

There being a multitude of possibilities, we have to confine ourselves to a very small fraction of them. If we do not confine the space of the solution, there are innumerable variants children can suggest, including looking for certain colors, forms, designs, or similar things (or combinations thereof) that do not already exist. We will confine ourselves now to all vertical-striped flags of Europe without any decoration.

One first has to make an inventory of existing flags—in this case, flags consisting of three cross-stripes (each a different color), without decoration, using only the given colors:

Belgium: black, yellow, red
France: blue, white, red
Ireland: green, white, yellow
Italy: green, white, red
Romania: blue, yellow, red
The total number of colors is 6.

How many possibilities are there for combining 3 colors? There are exactly 6 possibilities for combining 3 different colors (*A*, *B*, and *C*) in the way required: *ABC*, *ACB*, *BAC*, *BCA*, *CAB*, and *CBA*.

How many possibilities are there for selecting 3 different colors out of 6 different colors (*A*, *B*, *C*, *D*, *E*, and *F*)? *ABC*, *ABD*, *ABE*, *ABF*, *ACD*, *ACE*, *ACF*, *ADE*, *ADF*, *AEF*, *BCD*, *BCE*, *BCF*, *BDE*, *BDF*, *BEF*, *CDE*, *CDF*, *CEF*, and *DEF*—20 possibilities. (In general, we can deduce that there are $\binom{n}{k}$ possibilities to choose k elements, $k \leq n$, from n elements.)

A different order creates a new possibility. Therefore, our hero can make 6 × 20 = 120 suggestions for a new flag. Of course, 5 of these are existing flags, leaving 115 possibilities. So, if the ruler demands not more than 115 suggestions for a flag with the above conditions, our hero is saved. If he demands more, either the hero will have to die or we will have to outsmart the ruler.

There are innumerable possibilities for the creative student and teacher, by changing or extending the story.

Example 5

GAME
COMB
NUMB

Let us imagine a chessboard, where there is a knight (N) on the bottom left square. How can this knight reach the top row from there without moving backwards? What are the possibilities?

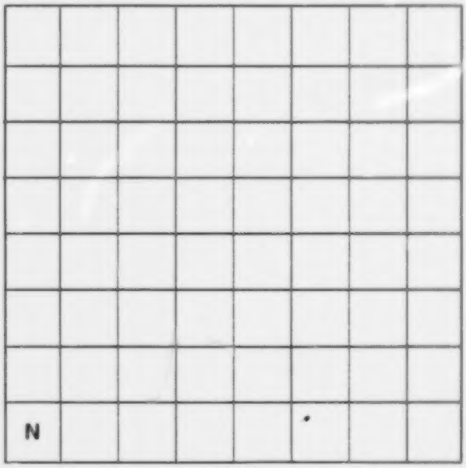


Figure 18.7

Again, the principles for finding the solution are to account systematically ⑪
for all possibilities and to solve it graphically, by making a picture or a table. ⑥

Solution

The knight has to bridge 7 rows.

One first has to appreciate the fact that the knight has generally two possibilities of progression: namely, to cross one or two rows of the chessboard.

Let us call the moves that cross one row "short moves" and the moves that cross two rows "long moves." Then we get the following equation: $7 = 2(a + b)$, with a the number of long moves and b the number of short moves.

As these are natural numbers, b has to be odd and less than or equal to 7. Otherwise, the sum could not be odd, because the sum of two even numbers is an even number. Hence, it follows that b can only be 1, 3, 5, or 7, and a can be 3, 2, 1, or 0.

The moves are classified as "short" or "long" only in the upward direction—moves to the right or left are not treated separately. Therefore, there are only the following possibilities for the knight in the left bottom corner to reach the top row:

1. 1 short move and 3 long moves
2. 3 short moves and 2 long moves
3. 5 short moves and one long move
4. 7 short moves

Therefore, for possibility 1, we get the following combinations:

- short, long, long, long
- long, short, long, long
- long, long, short, long
- long, long, long, short

So "short" can be the first, second, third, or fourth move.

For (2), there are 10 possibilities. For instance, the two long moves can come one after the other (4 possibilities: as first and second moves, second and third moves, third and fourth moves, or fourth and fifth moves); with an interval of one move (3 possibilities: as first and third moves, second and fourth moves, or third and fifth moves); with an interval of two moves (2 possibilities: as first and fourth moves or second and fifth moves); or with an interval of three moves (as first and fifth moves). The remaining moves have to be the short ones.

For (3), one can find the possibilities by considering the following: The "long" can take any position from the first to the sixth. So there are 6 possibilities for the order of the moves.

For (4), there is obviously only 1 possibility. In general, the number of combinations of a long moves and b short moves is given

$$\frac{(a+b)!}{a!\,b!}.$$

All 21 different possibilities, according to the order and number of long and short moves of the knight, lead right to the aim. The knight comes from the bottom left corner to the top row of the chessboard.

The task can be variously extended or generalized, or analogous problems can be created. For instance, all concrete possibilities of the positions of the knight's moves could be found. One could also abandon the condition of not going backwards. Then, with the restriction that no square may be touched twice, there would be at most 63 possible moves for the knight.

It is really a demanding brain teaser for students interested in chess mathematics to exhaust this maximal number of moves and touch all squares at most once.

To solve this task in general, one needs some knowledge of grid-point theory (number theory—divisibility, discrete geometry—reflection). This task is also called Euler's knight's-jump problem.

The interested teacher can find further chess-related mathematical problems in Gik's (1986).

Outlook

What student isn't intrigued once he learns that famous mathematicians, engineers, musicians, and artists of all fields approached and solved new problems creatively and playfully? Their joy in playing turned out to be the secret of their success.

Thanks to innovative teaching and learning strategies, students can recognize that levity helps destroy the usual mental barriers. Playing and academic performance are not exclusive of one another, as it is unfortunately often supposed. Students' joy in playful thought can be shaped in diversified and productive ways. The playful wording of mathematical problems and the discovery of mathematical principles in games can promote better concentration in students, as well as long-lasting interests.

Children who are allowed to become absorbed in their playing are capable of a better intellectual performance than children who are drilled and pressured. Motivation by playing does not lessen the incentive and the desire to finish even difficult tasks, but rather promotes these qualities.

The overlapping teaching strategies recommended in this chapter fall under the pedagogical term "ludification." Ludification aims at the development, creation, and improvement of the motivation for mathematical tasks by using game elements and game strategies or, conversely, the discovery of mathematical problems through games.

Only by succeeding in this can teachers lead their students into creative mathematical problem solving.

References

Bönsch, E. (1987). *Schachlehre.* Berlin: Sportverlag.

Gik, J. (1986). *Schach und Mathematik.* Berlin: Urania-Verlag.

Kauke, M. (1992). *Spielintelligenz: Spielend lernen—Spielend lehren?* Heidelberg: Spektrum Akademischer Verlag.

Klaus, G. (1969). Was und wem nützt Schach? *Schach,* 8.

Posamentier, A. S., & Stepelman, J. (1995). *Teaching secondary school mathematics: Techniques and enrichment units* (4th ed). Columbus, OH: Macmillan/Merrill.

19

Cooperative Learning Approaches to Mathematical Problem Solving

HOPE J. HARTMAN

This chapter gives an overview of rationales for and methods of using cooperative learning to teach mathematical problem solving. Cooperative learning involves students working together in a teaching-learning situation. There are at least three basic forms of cooperative learning: tutoring (peer or cross-age), in which one student teaches another; pairs, who work and learn with each other; and small groups of students teaching and learning together. First this chapter focuses on teachers' and researchers' conclusions about cooperative learning as an effective method of improving mathematical problem solving. Next it describes several methods of using cooperative learning to teach mathematical problem solving. Then the chapter concentrates on the teacher's role in cooperative learning and concludes with some pitfalls involved in cooperative learning.

Why Use Cooperative Learning in Mathematics?

Sometimes the best way to learn something is to teach it. Teaching requires considerable depth of knowledge, memory of important concepts, understanding, organization, and skills. Cooperative learning provides opportunities for students to teach each other. When students explain and teach concepts to each other, retention of these concepts improves (Webb, 1985). Explaining also helps students connect their prior knowledge with new information (Pressley et al., 1992). There are many good reasons for using cooperative learning in mathematics. Crabill (1990) came up with at least 16 advantages to using cooperative learning in high school classrooms. Contrary to some current speculation, cooperative learning in mathematics is more than an educational fad. It has been used successfully in the Netherlands since the early 1970s (Terwel, 1990), and word is getting out internationally that this is a powerful instructional method for developing mathematical problem-solving skills.

It is useful for professional educators to have a repertoire of teaching strategies, so they are able to be flexible and shift as situations require; variety also helps prevent boredom. Many math classes have students of widely ranging achievement levels, and there is also an increasing amount of ethnic and linguistic diversity in U.S. classrooms, where cooperative learning has been demonstrated to be an especially effective method of teaching. Cooperative learning in mathematics can be done at any age, often with teachers' existing curricular materials, and has been successful in improving mathematics achievement from elementary grades through college (Frankenstein, 1986).

Students can learn to solve mathematical problems and can improve their understanding of mathematics concepts without lecture or other forms of direct instruction. Through studying worked-out examples of factorization and learning by doing their own problems, students learn to recognize when the procedures are applicable (Zhu & Simon, 1987). Cooperative learning is another alternative to direct instruction and lecture, requiring students to be more actively involved in the learning process. The social nature of cooperative learning and the relative emphasis on student control over instruction make it fun and highly motivating for most students. Because the teacher is more of a manager of instruction than a transmitter of information, cooperative learning promotes student self-regulation as students make, detect, and correct their own and each other's errors.

The National Council of Teachers of Mathematics (NCTM) supports the use of cooperative learning for teaching mathematics because it enables students to discuss and learn to listen to each others' ideas, ask questions, make mistakes, and offer constructive criticism. Students help each other connect new information with what they already know and discover their own meaning as they explore ideas that arise in groups (Artzt & Newman, 1990).

Mathematical Thinking and Achievement

By comparing problem-solving processes and outcomes with their peers, students can learn to differentiate between appropriate and inappropriate strategies to use in different cases. Learning this from other students rather than from the teacher can demystify the problem-solving process and reduce math anxiety. Anxiety and misconceptions about mathematics can contribute to poor mathematics self-concepts and can inhibit success in mathematical problem solving (Gourgey, 1992). Too often, students feel dependent on the teacher's expertise for learning mathematics. In order to think mathematically—one component of which is problem solving—students also need to develop their metacognitive skills of planning, monitoring, and evaluating so they relate these skills to their problem solving (Schoenfeld, 1989). Cooperative learning is an especially effective method of activating metacognitive aspects of problem solving (Artzt & Armour-Thomas, 1992).

Studies comparing cooperative learning with competitive and individualistic learning have demonstrated that cooperative learning promotes higher achievement in mathematics than the other two methods. Not only do students solve math problems more successfully and learn and retain mathematical concepts, but cooperative learning also results in more use of higher-level thinking, more frequent discovery, more new ideas and solution strategies, and more transfer of what is learned in groups about problem solving to individual problem-solving situations. These benefits are a result of student internalization of math concepts and problem solving through their discussions and explanations of problem-solving strategies and approaches with their peers (Johnson & Johnson, 1990). Giving mathematical explanations requires deeper understanding than just putting an answer on a worksheet (Artzt & Newman, 1990).

① ⑤ There are numerous effective strategies students can learn (such as working backwards and trying extreme values) to help them become effective problem solvers. A major benefit of cooperative learning is that it allows students to compare the strategies they use and discuss the advantages and disadvantages of using them for different problems. When structuring groupwork around instruction in such problem-solving techniques, make sure groups of students ask and answer key questions about their strategy use: What is the strategy? Why is it a good strategy to use in this situation? How will it be applied? What other strategies could be used? What are the advantages and disadvantages of each? Through dialogue with other students about the use of problem-solving strategies, misconceptions can be clarified and strategy use improved.

Students can work cooperatively on mathematically significant, interesting, and complex tasks. Through working with others they can enhance their ability to communicate about mathematics, to understand it, and to think critically about it (Good, Reys, Grouws, & Mulryan, 1989/1990). Discussing problems helps students become aware of what they know and do not know and what they understand and do not understand. This awareness leads, in turn, to control over the problem-solving process as students begin to see the need for specific problem-solving strategies and the limitations of other approaches. When students find that their approaches and outcomes for the same problem differ, this discrepancy can stimulate reorganization and development of their thinking to a new and higher level. ③

Scaffolding

Cooperative learning can be used as a strategy for scaffolding instruction from teacher direction and control to student self-regulation. Scaffolding means providing temporary support (models, cues, prompts, hints, partial solutions) to students to bridge the gap between what students can do on their own and what students can do with guidance from others. The goal of scaffolding is for students to become independent, self-regulating, self-sufficient, and less teacher dependent. Scaffolding is an especially effective teaching approach for developing higher-level cognitive strategies, such as those involved in problem solving. Based on his experiences with small-group problem solving, Schoenfeld (1989) believes it is valuable for giving teachers a chance to give students support and assist them while they are actively engaged in the problem-solving process. Scaffolding often involves the following basic components:

1. Present the new cognitive strategies.
2. Regulate difficulty during guided practice.
3. Provide varying contexts for student practice.
4. Provide feedback.
5. Increase student responsibility.
6. Provide independent practice. (Rosenshine & Meister, 1992)

③ Many educators believe that to learn mathematics, students must construct and reconstruct mathematical concepts, relationships, and procedures in their own minds and within the contexts of meaningful problem-solving situations. Teachers can start with think-aloud modeling and direct instruction, then shift to cooperative-learning formats for extensive and varied practice with feedback. Students can then guide each other to effective problem solving. As students generate and evaluate alternative approaches, they learn that there are different ways of solving problems and that some ways are better than others; they internalize feedback and gradually learn to detect and correct their own errors, enabling them to solve problems on their own.

Personal and Social Aspects of Cooperative Learning

Problem solving is usually much more fun and exciting to students when they work together. The social context and active involvement make it more intrinsically motivating to learn mathematics. Research has shown that cooperative learning increases confidence in students' mathematical abilities and improves self-esteem and feelings of self-efficacy in approaching math problems. These effects increase the likelihood of students—especially for females and minorities—choosing careers in math and science (Jackson, 1989). Other research has documented the effects of cooperative learning on improving interpersonal relations with students of different ethnicity and cultural backgrounds (Johnson & Johnson, 1990).

Students using Reciprocal Peer Tutoring (RPT) improved not only their math achievement but also their adjustment to school. School adjustment included student self-perception measures of behavioral conduct, scholastic competence, social acceptance, and global self-worth, as well as teacher-reported conduct (Fantuzzo, Alperin King, & Rio Heller, 1992).

Specific attitudinal objectives targeted by cooperative learning include positive mathematics attitudes, confidence in one's mathematical thinking, willingness to take risks and try various strategies, accepting frustration, persevering when solving difficult problems, and attributing failure to not yet using the right strategy yet rather than to lack of competence (Johnson & Johnson, 1990). Studies have documented that cooperative learning in math classes improves students' attitudes toward math as a subject and toward math instruction. Students feel more confident in their ability to do math and are therefore less anxious about it. Cooperative learning enables students to get and receive help in a relatively nonthreatening context (Artzt & Newman, 1990).

How Does Cooperative Learning in Mathematical Problem Solving Work?

The following is a simple cooperative learning segment of a lesson from a senior honors math class. The class was studying related-rate word problems. The problem was:

> A boat is pulled in to a dock by means of a rope with one end attached to the bow of the boat, the other end passing through a ring attached to the dock at a point 4 feet higher than the bow of the boat. If the rope is pulled in at the rate of 2 feet/second, how fast is the boat approaching the dock when 10 feet of rope are out?

ALG

Students were assigned to heterogeneous groups of three or four students to work on the problem. Their task was (a) to individually generate questions for solving the problem, (b) to share their questions with the group, (c) to decide as a group on the best questions for this problem, (d) to solve the problem individually using the group's questions, and (e) to share and compare individual solutions and explain how they were obtained from applying the questions selected. Students were taught to use self-questioning as a strategy for thinking through the problem-solving process. Initially, the students found it strange to be asked to write questions for a mathematics class. They learned how to use questioning to help them plan, monitor, and evaluate problem solving. The teacher modeled examples of questions by thinking aloud. Then students generated and used their own questions. The three groups generated questions for this problem as shown here:

Group 1

⑥ Student 1: What should the diagram look like? Where do the values belong? What do I want to find?

Student 2: Where do I start? How do I find the desired answer? Where do the numbers belong in the formula? Which number goes to which part?

Student 3: What does the diagram look like? What variables should I use? Where does the 2 feet/second go? What derivatives do I have to find?

Student 4: How do I draw a picture to represent what the problem says? What parts of the diagram get labeled? What is the unknown? What equation do I use to get the derivative?

This group discussed their questions and made the following list to use when solving the problem:

1. What should the diagram look like?
2. How should it be labeled?
3. What do we have to find?
4. What equation do we use to find the derivative?

Group 2

Student 1: What does this look like? What formula do I use? How do I approach it? What do I want to find?

Student 2: Are they on the water? How long is the rope? Is this a controlled area with no waves and no current? How heavy is the boat?

Student 3: What do I have to find? What speed, velocity, and rate will help me solve the problem? Why can't I figure this out? How do I differentiate the problem with respect to time?

Student 4: What am I given? What must I find?

After discussing their questions, this group made the following list:

1. What is given and what must be found?
2. What variables must be considered and how?
3. How do we approach this?

Group 3

Student 1: What do I find? What does the diagram look like? How do I draw this? Is it similar to something we've done? What is the equation? Is my algebra correct? Am I using the right formula? Is the diagram correct?

Student 2: What type of diagram will this be? What are the dimensions of the diagram? What is the rate of the rope being pulled in? What is the problem looking for? What is given? What formula will be needed?

Student 3: Is the rate 2 feet/second horizontal or vertical? Where does the equation come from? Is 10 feet used right now or after the derivative? How is the Pythagorean theorem used if the hypotenuse isn't there?

This group's discussion led to the following set of questions:

1. What is given?
2. What should the diagram look like?
3. Is the diagram correct?
4. Is this similar to something we've done before?
5. What formula will be needed?
6. Is this the right formula?
7. What do we do with the speed and distance of the rope?
8. Is the algebra right?

While the groups worked on their questions and used them to solve the problem, the teacher walked around to watch and listen to each group to make sure they were on track and making reasonable progress. As she checked up on each group she saw that some students still could not solve the problem. She checked the individual and group lists of questions and saw that they were incomplete, so she decided to have the groups share their questions, evaluate them as a class, and come up with a composite list. She guided the discussion to make sure the class generated questions for all three phases of the problem-solving process (planning, monitoring, and evaluating). The following composite list emerged:

Planning

1. Does this problem resemble a problem already done?
2. How should I diagram this problem?
3. What do I have to find?
4. What equation must I differentiate?

Monitoring

1. Is my algebra correct?
2. Am I using the correct formula?
3. Is my diagram labeled correctly?

Evaluating

1. Does the answer make sense?
2. Did I find what I was supposed to find?
3. How can I check my answer?

Students then returned to solving the problem with the new set of questions. Individuals within each group shared their answers with each other, decided on the correct answer, and raised their hands to let the teacher know when they were finished so she could check their solutions. She randomly asked students in the groups to explain their solutions to make sure everyone in the group understood the problem and solution process. Then she had students who had solved the problem help those who had difficulty. At the end of the lesson the class looked at how the questions related to each part of the problem solution. Constructing, comparing, discussing, and evaluating problem-solving questions individually, in small groups, and with the entire class enriched students' understanding of what questions and strategies were best suited for the particular problem. Some of the students said that in the past, they had been so concerned with getting the right answer that they had never given as much thought to the thinking process.

③ When each student has just his or her own knowledge, thoughts, and questions, the perspective on problem solving is much more narrow and shallow. Mathematicians frequently discuss their solution strategies and outcomes with others. They know that others can sometimes detect limitations and suggest alternative approaches to, and applications of, problem solutions. By discussing problem solving with others, students learn to think more like mathematicians.

The Math Solution

The Math Solution (Burns, 1990) is an elementary school in-service program for teaching mathematics as a tool for solving problems and as a way of thinking. It is based on the assumption that for math learning to occur, students need maturity, physical experience, and social interaction. Communication between students should be encouraged so that all children get experience explaining and clarifying their thinking and can move from subjective to more objective views. This program recommends teaching mathematical problem solving in heterogeneous *groups of four*, seated together. Three rules govern groupwork:

1. Students are responsible for their own work and behavior.
2. Students must be willing to help any group member who asks.
3. Students may ask the teacher for help only when everyone in the group has the same question.

Encouragement, practice, and discussion are needed for students to learn to work together successfully.

Three stages of instruction make up a Math Solution lesson. Lessons may vary in length, lasting from a class period to a week or longer. The stages are introduction, exploration, and summarizing. Whole-class instruction occurs during the introduction stage, when teachers (a) review or present concepts that are needed, (b) pose a similar or smaller problem or part of the problem for students to try, (c) present the problem to be solved, and (d) discuss to make sure students understand what to do. Cooperative learning begins with the exploration stage. As students work on the problem, the teacher (a) observes student interaction, ④ listening to each group's strategies, procedures, and ideas; (b) assists only when needed (if hands are raised or the group is not working); and (c) provides an extension activity for groups that finish early. The teacher may pose several problems at once so groups can continue working at their own pace. The third stage, summarizing, has three goals. First, groups share their problem-solving processes, procedures, and strategies. Students critically evaluate their approaches and consider alternatives for future application. Second, groups present solutions, showing their work whenever possible. During this stage, it is recommended to ask, How did you decide if your findings made sense? How can you check the solution ⑱ (Burns, 1990, p. 32)? Third, generalize from the solutions. During this stage, students abstract what they did in solving a particular problem to think about how they would approach related problems. Questions to ask include, Are there patterns or relationships you can see from your solution? Does the problem remind ② you of another problem you have solved? How are they alike or different (Burns, 1990, p. 32)?

Finding Out/Descubrimiento

The Finding Out/Descubrimiento mathematics and science curriculum (Lotan & Benton, 1990) was developed specifically for bilingual elementary school students. It is designed to develop thinking skills and improve the academic and linguistic performance of children in culturally and socially heterogeneous classrooms. Since 1979, research has demonstrated this curriculum to be successful in promoting student achievement in mathematics and science. The key features of the approach are differentiated tasks, delegation of authority, student interaction, and the treatment of status. Proponents of the method argue that "the development of thinking skills requires increased amounts of task-related interaction, through which students have the opportunity to develop problem-solving strategies" (Lotan & Benton, 1990, p. 64). Students work together, continually communicating about their hands-on work at learning centers, where they learn to use problem-solving strategies and explore the world around them. Complex, multiple-ability tasks are carefully selected for cooperative, discovery-oriented learning. Two instructional features are built into this model to address the inequalities of heterogeneous groups: multiple-ability treatment and assigning competence. The former involves a direct statement by the teacher that many abilities are needed to perform each task. No one is good at absolutely all of them, and everyone is good at some of them. The teacher emphasizes this point repeatedly and points out specific skills and abilities needed for various tasks. Teachers assign competence to low-status students to help equalize their interaction, because low-status students are often perceived as unable or unwilling to help others and are generally less influential. Observing and recognizing low-status students being competent is the essence of this technique. The teacher looks for low-status students making valuable contributions to the group and then publicly identifies the particular skill reflected by the contribution. For example, a low-achieving student may be excellent at cutting out the shapes needed for a group task, which reflects good visual, spatial, and motor skills. This person gains status as the group's official cutter. As a result of public recognition, low-status students often raise their expectations for their own performance.

Real Maths

The Real Maths curriculum (Terwel, 1990) for students ages 12 to 16 is based on the theory of Van Hiele, an internationally known Dutch mathematics teacher and researcher. Van Hiele's level theory identifies three levels:

Zero or perceptive level: Students look at the whole problem without analyzing the parts.

First or descriptive level: Students describe the parts and their characteristics and have an intuitive understanding, but there is no reflection on fundamental ideas.

Second or theoretical level: Students' intuitive concepts are formulated more explicitly, and students reflect on concepts and on the relationships between problem parts and the whole.

Van Hiele recommends a five-stage teaching-learning process:

1. Information: Students get materials (e.g., objects, graphs, papers) to use in exercises.
2. Structured orientation: Students are assigned specific tasks. Each task is designed to teach students one characteristic of the material they are using.

3. Explicating: Students describe the characteristic verbally.
4. Free orientation: Students are given general tasks that require them to find their own way in a network of relations.
5. Integration: Students reflect on different solutions, explore relationships between them, and formulate laws of a new and higher-level structure. ③

The Dutch were pioneers in using cooperative learning in mathematics. Real-life settings are essential features of all mathematics problems in this approach. The lesson-design model for this curriculum consists of three stages:

1. Introduction: Working with the whole class, the teacher introduces the problem, explores aspects of it, and may give hints about the solution and place it in an everyday-life context.
2. Groupwork: Students work in groups as the teacher observes and manages their cooperative problem solving. When necessary, the teacher deals with individual problems.
3. Reflection and evaluation: Students discuss several topics related to their group process and results. The discussion involves identifying all the different solutions and strategies groups used to solve the problem, questioning (by the teacher) to explore other possible solution strategies, reformulating and summarizing solutions, and generalizing about solutions. ⑱

Sample lessons involve using newspaper ads, deciding which video rental shop to become a member of, and selling badges as a small business enterprise. The Real Maths curriculum has accompanying teacher's manuals as resources for using real-life situations for mathematical problems and for teaching with small, heterogeneous groups. It also has student materials and videotapes.

Team Accelerated Instruction

Team Accelerated Instruction, or TAI (Slavin, 1990), involves specially developed curricular materials for cooperative learning in mathematics. Formerly called "Team Assisted Individualization," the major components of this approach are teams, placement tests, curriculum materials, team study, team scores and recognition, teaching groups, fact tests, and whole-class units. TAI has been used at elementary, middle, and high school levels. This approach combines individualized instruction with cooperative learning and individual accountability with group rewards. Slavin suggests that math, in particular, requires individualization, because students often show a wide range of individual differences in ability or achievement, learning in mathematics being so dependent on prior knowledge and skills. TAI was developed to address the high level of student heterogeneity in math classes. A middle school teacher who used TAI commented, "The teacher has flexibility to vary from group to group or individual to individual. TAI lets students process a lot of the paperwork that ties a teacher down. Students check each other's work as they progress through the units, . . . [providing the] immediate feedback students need [as] it identifies problems that often can be handled in the group or answered by the teacher if further help is needed" (Slavin, 1990, p. 81). A fifth-grade teacher reported that individual accountability and group rewards are especially effective for students who dislike mathematics. TAI enables students to work at their own level and achieve success, thereby eliminating frustration and boredom. Student success is translated into team points, which motivates students to do well. The individual student gets self-satisfaction; the group gets a reward and positive reinforcement from the teacher. A high school special-education teacher who used TAI noted that working toward a team score led students to encourage each other to complete class assignments and resulted in a faster rate of mastery than did individualized mathematics instruction. Mastery of 80% is required to move to the next level, so the quality of work remains high.

Tutoring and Pair Methods

Reciprocal Peer Tutoring

Reciprocal peer tutoring, or RPT (Fantuzzo et al., 1992), is for elementary students learning mathematical computation. In this approach, two or more students work together cooperatively and follow a structured format in which they teach, prompt, monitor, and evaluate each other. Students alternate between teacher and student roles and engage in peer teaching, peer choice of rewards, and peer management. Variations of RPT have been used successfully in urban, low-income public schools with minority and white elementary students. Students are trained to use the procedures in two or three sessions of 45 minutes each. Modeling the procedures and using instructional prompts enable students to use these procedures without assistance. Training includes discussing with students the value of teamwork, partnership, and cooperation. A unique aspect of this approach is the emphasis on student involvement in the reward structure. The use of group rewards has been found to improve conduct in the classroom. It is the combination of these components (peer teaching, peer choice of rewards, and peer management) that produces greater academic and motivational gain than using them in isolation.

The 6PQ Method of Discovery Learning

The 6PQ Method (Chemeketa Community College, 1986) is a six-step question process designed to guide student thinking in the right direction while avoiding telling answers and lecturing. Its purpose is for students to be active, independent thinkers and learners. Each of the six question types starts with the letter *P*. The tutor frequently paraphrases students' responses to questions within and between steps. This method can be used as a tutoring technique in a peer or cross-age tutoring program or for classwide peer tutoring in the regular classroom. Pairs of students can take turns solving problems, either alternating roles as tutor and student or with stronger students tutoring weaker students. The entire class gets trained in this learning method so that all understand what happens and why. This method not only can help students become active learners and independent problem solvers, it also develops students' questioning, communication, and leadership skills while simultaneously building their self-esteem and mathematics self-concepts. The six steps to use for problem solving include

1. Preface: Establish rapport and identify the problem to be solved.
2. Pace: Determine what the student already knows about this problem and how to solve it. If the student knows how to solve the problem, observe the solution process and go to step 6.
3. Probe: Investigate the limits of the student's knowledge by seeking more detailed information. If the student solves the problem, move to step 6.
4. Prod: If the student still has not solved the problem, ask him or her to make an educated guess about how to solve it, or ask how the student feels (rather than thinks) it might be done. If the students solves the problem, go to step 6. ⑦
5. Prompt: If the student is stuck and all prior attempts have failed, give a hint or suggestion, or show the student how to start the process. If this works, move to step 6.
6. Process: Have the student apply the same problem-solving process to a similar problem. ⑱

The tutor should paraphrase frequently but not continuously. At the end of questioning, ask the student to summarize what was learned. See the dialogue below for an example.[1]

T = tutor, S = student

1. Preface

T: Hi Celeste! I'm happy we get to work together again.
S: Me too Alicia! I need your help. I'm stuck on one of our homework problems.
T: So you want me to help you with a specific problem? [paraphrase]
S: Yeah, I figured out some of it, but now I'm stuck.
T: What problem are you working on?
S: A factoring problem, number 4: Factor $6x^3 + 12x^2 + 6x$. ALG
T: So you only need help solving part of the problem? [paraphrase]
S: Yeah. Do you know how to do it?
T: Yes. I'm pretty good at factoring.

2. Pace

T: What do you know about this problem so far?
S: I think this is called a polynomial.
T: Yes, that's right.
S: And I know I have to pull out what the parts have in common. So first I pulled out a six from each part. That left me with $6(x^3 + 2x^2 + x)$. Is that right?
T: That's good so far, what else do you know?
S: Next I saw that all parts have an x, so I pulled that out.
T: What did that leave you with?
S: I got $6x(x^2 + 2x + 1)$. Am I OK so far?
T: Yeah, you're doing fine. What comes next?
S: I don't know. This is where I get stuck.
T: OK. So you know you can factor out $6x$ from the polynomial, but that's where you run into trouble, is that right?
S: Yes.

3. Probe

T: Can you tell me more about the trinomial you have left, $x^2 + 2x + 1$?
S: I remember something about two sets of parentheses, but I can't remember what numbers to put in them.
T: Do you know the name for the trinomial?
S: What do you mean by a name?
T: There's a specific equation the trinomial is associated with.
S: Is that the quadratic equation?
T: That's right. What else can you tell me?
S: Nothing. Nothing at all.
T: OK, so you know that you arrange numbers in two sets of parentheses and that somehow this is connected with the quadratic equation, but that's it. [paraphrase]
S: Yes.
T: How are the parentheses set up?
S: They are next to each other.
T: And what does that mean?
S: That you multiply what's inside each of the parentheses. Oh! So I can put an x inside each parenthesis and multiply them to get x^2. Is that right?
T: Yes, good! When you multiply the xs in the parentheses, it results in the first term of the quadratic equation.
S: I guess I knew more than I thought!

⑦ 4. Prod

T: OK, now look at the problem again and make an educated guess about how to fill in the rest of the quantities.

S: Hum . . . I'm getting a little clearer about it. I've got to put the numbers and xs into the parentheses so that when they are multiplied together I get the trinomial $x^2 + 2x + 1$. But I don't know what to do to get it—what else to put where.

5. Prompt

T: I'll give you a little hint. Remember what function you used to get the x^2 from the xs in the parentheses?

S: I multiplied.

T: Right, so what function will you use to get the $2x + 1$?

S: I guess I have to multiply to get the last term—oh yeah, and then I add to get the one in the middle! I have to find numbers that when multiplied you get one and when added you get $2x$. So what would that be? Let's see . . . one times one equals one, and one plus one equals two. That must be it! Is it?

T: Sounds like you got it! So what does it all look like?

S: It must be $6x(x + 1)(x + 1)$.

T: Great, that's it! But can you take it one step more? Isn't there a simpler way of representing the information in the two sets of parentheses?

S: Oh, I see, $(x + 1)$ $(x + 1)$ can be changed to $(x + 1)^2$. So that means it ends up as $6x(x + 1)^2$. Is that it?

T: Alright, now you got it!

6. Process ⑱

T: How can you use what you learned to solve other problems?

S: Well I was stuck on homework problems 7, 24, and 30, because I think they are similar problems that I got stuck on the same general way.

T: Yes, they're similar. So you think that you can solve problems 7, 24, and 30 now based on what you have learned? [paraphrase]

S: I think so. Thanks for the help!

T: Will you summarize what you learned before you try those other problems?

Summary

S: I started with the problem: Factor $6x^3 + 12x^2 + 6x$. I saw that I could pull out the 6 as a common factor, and that I could also pull out the x, but then I got stuck. Now I know to arrange the numbers in parentheses side by side so when they're multiplied they give me the end terms in the trinomial and when added they give me the middle term.

In this method, telling or showing how to solve the problem is only used as a last resort. Through 6PQ, students learn that they know more than they think they do, which helps build their confidence and willingness to persist.

Pair Problem Solving

Thinker and listener pairs working on problems and rotating roles has become a popular way of helping students think about their own problem solving (Whimbey & Lochhead, 1982). Students take turns serving as thinkers (problem solvers) who externalize their thought processes by thinking aloud, while analytical listeners track and guide the problem-solving process as needed. It is a higher-level thinking (metacognitive), self-monitoring strategy that gives students feedback on what is understood and what is still unclear. It helps students identify what parts of a problem they understand and where they get stuck. This method makes problems more engaging, teaches communication skills, and fosters cooperation. It encourages skills of reflecting on beginning and later thoughts. Pair problem solving facilitates self-checking and evaluating. It encourages the formation of study and support groups and exposes teachers and students to various solution approaches (Heiman, Narode, Slomianko, & Lochhead, 1987a). By listening to his or her own thoughts, the student gains awareness and control over problem solving. Externalizing thoughts enables them to be seen from a fresh perspective.

How to Think Aloud: Problem Solver's Task

1. Translate your thoughts (ideas, images, etc.) into words and recite them aloud.
2. Verbalize aloud all the steps you go through when solving problems. No thought or step is too small, easy, obvious, or unimportant to verbalize.
3. Verbalize all the thinking you do before you start to solve the problem (e.g., what you are going to do, when, why, and how). Even second-guessing yourself is important to verbalize aloud—for example, "I think I should use that long, complicated formula we were using a couple weeks ago. What was it called, the quadratic equation? No, maybe not. Maybe I'm supposed to use the formula we did in class yesterday."
4. Verbalize all thoughts during problem solving—for example, "OK, I'm almost through with this division problem. Now that I have my answer, all I have to do is multiply to check and see if my answer is right" (Hartman, 1992).

How to Listen Analytically: Listener's Task

1. Think along with the problem solver. Follow every step and make sure you understand every step. If not, ask a question. Have the problem solver identify and define important terms, variables, rules, procedures, and so forth. Make sure the problem solver vocalizes all the steps and does all the work. If the problem solver skips over a step without thinking aloud, ask her or him to explain the missing thought.
2. Do not work on the problem independently. Listen to and work along with the problem solver.

3. Never let the problem solver get ahead of you. Whenever necessary, ask the problem solver to wait so you can check a procedure or computation and catch up. If the problem solver is working too fast, slow her or him down so you can follow carefully, analytically, and accurately.
4. Check the problem solver at every step. Don't wait for the answer. Check everything—each computation, diagram, procedure, and so forth. In the back of your mind, constantly ask yourself, "Is that right? Did I check that?"
5. If you find an error, avoid correcting it. Point it out and try to get the problem solver to self-correct. If he or she gets stuck, ask questions to guide thinking in the right direction. If necessary, give suggestions, hints, or partial answers. Give the answer as a last resort. Let the problem solver know that you are not trying to be difficult but are trying to help him or her become an independent problem solver (Whimbey & Lochhead, 1982).

To make sure listeners really do their job, teachers should periodically ask listeners to summarize the steps the problem solvers used. The following activities were suggested by Larcombe to help student pairs in elementary grades or remedial high school students learn to externalize their mathematical thought processes:

1. Students take turns describing the rules they use.
2. Students describe to each other how the parts fit together when doing a construction task.
3. Working with concrete objects at first, students can describe operations used when calculating. ⑥
4. One student must guess an object, mathematical representation, or graphic based on another student's description. (Heiman et al., 1987a)

I DREAM of A

I DREAM of A (Hartman, 1992) is a technique for developing mathematical problem-solving skills that synthesizes thinking aloud and questioning. It is a math-specific adaptation of Bransford and Stein's (1984) *IDEAL Problem Solver*. Each letter in the acronym stands for a component of the problem-solving process. These components involve metacognitive skills for planning, monitoring, and evaluating the problem-solving process. The first four letters are all planning steps, which may be performed in different sequences; the I DREAM of A approach is not intended to be a rigid formula. It should be adapted to the individual problem solver's needs in each problem-solving situation. The teacher can serve as an expert model, demonstrating how to use I DREAM of A by thinking aloud and self-questioning while solving problems. Then two students work together, as in the pair problem-solving method described above. One student is the questioner, guiding the problem-solving process by questioning the other student and by having the problem solver think aloud periodically. The problem solver answers questions and thinks aloud, as prompted by the questioner. Questions are asked for each of seven components I DREAM of A. Although most questions focus on knowledge and strategies needed to solve a problem, the questioner occasionally asks about the problem solver's feelings to establish and maintain a positive attitude. The questioner decides what questions to ask, when to ask the problem solver to think aloud, and when to ask about the problem solver's attitudes. An illustration of the model follows.

⑭ *I • Identify*

What is given/to be found? How could you state the problem in your own words? What do you know about this type of problem? Does anything seem confusing? What information is relevant? What isn't? How do you feel about your ability to solve this problem?

⑥ *D • Diagram*

How would you sketch or draw a picture of that problem? What symbols or notation should be in your diagram? Are you sure you have all the important parts? Are you sure they are in accurate relationship to each other? How does your diagram help you identify and define the problem?

④ *R • Recall*

Have you solved problems like this before? How? Do you have an example of how to solve this type of problem in your text or notes? What definitions, rules, concepts, procedures, equations, and such might you need to solve this problem? Why? When would you use them? Is there anything you should review? Are there any mistakes you tend to make on these problems?

① *E • Explore*

⑤ Think out loud about how you might solve this problem. What approaches could you use? Are there any general strategies, such as working backwards or trying extreme values, that might be useful? Which is best? Why? What would you estimate the answer to be? What difficulties might arise while solving the problem? How could you break the problem into parts? What would you do first, second, . . . ?

A • Apply

What is your plan for solving this problem? Think out loud as you apply your problem-solving plan.

M • Monitor

How are you doing so far? Is your approach leading you where you need to go? How can you tell? Should you try another approach? How have you been checking to make sure you haven't forgotten anything and haven't made any careless mistakes? Are you keeping a positive outlook? Does everything make sense?

A • Assess ⑱

Are you sure you have answered the problem completely? Looking back at the problem statement, does you answer make sense? How can you verify your answer? What did you learn from solving this problem that could help you in the future?

In a peer or cross-age tutoring situation, the tutor is the questioner and the person being tutored is the problem solver. In pair learning or cooperative learning, students take turns performing roles of questioner and problem solver. To become independent problem solvers, students can learn to ask themselves such questions when working on their own and internalize the components of the problem-solving process. The sample questions are presented to stimulate thinking about appropriate types of questions to ask. Each situation will require somewhat different questions. Research shows that student-generated self-questions are more effective than teacher-imposed questions. Teacher-imposed questions are, however, good building blocks or scaffolds which temporarily guide students through problem solving until students can independently self-question, maintain a positive attitude, and guide their own problem solving.

The Teacher's Role in Cooperative Problem Solving

One of the key features of cooperative learning is that the teacher delegates considerable responsibility to the students. Both teachers and students must adopt roles that are quite different from whole-class instruction. "The essence of good problem solving is self-correction. . . . Teachers should become facilitators of learning, not sole dispensers of truth, [and] must relinquish the safe seat of authority and step into the classroom . . ." (Heiman et al., 1987b, p. 8).

An important task in making cooperative learning successful is attitude change. The teacher's own attitude has to shift from teacher as transmitter of knowledge, center of attention, and authority to teacher as manager and facilitator of learning. Attitudes toward noise in the classroom sometimes must be changed. Many people equate a noisy classroom with disruption and chaos. In cooperative learning, noise in the classroom can be a sign of high-level thinking and learning. Several student attitudes must change so students show interest in finding solutions, confidence in taking risks and trying various strategies, willingness to be wrong, acceptance of frustrations, perseverance when solutions are not immediate, and understanding the difference between not yet having found an answer yet and not knowing it. Students must shift from depending on teachers for the answers to becoming independent thinkers and learners (Johnson & Johnson, 1990). Teachers as role models must emphasize to students the importance of being problem solvers, active learners, and seekers who are willing to take risks and make errors. Mistakes should be treated as learning opportunities, and students must understand that important aspects of problem solving lie beyond the correct answer. Teachers and students alike must learn to value the process, not just its products.

Erickson (1989) devised a list of nine steps for teachers to follow in implementing cooperative learning:

1. Ensure a successful experience the first time.
2. Decide what to watch for.
3. Decide on a grouping strategy.
4. Prepare the materials.
5. Prepare yourself.
6. Explain the rules and expected behaviors.
7. Do it!
8. Debrief the class.
9. Debrief yourself.

Planning for Cooperative Learning in Mathematics

Training Students for Cooperation

Although many everyday life activities require cooperation, seldom are students taught how to work together cooperatively. For cooperative learning to work effectively, some educators recommend giving specific training in skills needed for cooperative learning. These skills include observing carefully, reasoning, asking key questions, being supportive and helpful to others, explaining clearly, thinking visually, reasoning spatially, recording data, exploring new solution strategies, understanding the problem, being persistent, and using ideas of other students (Lotan & Benton, 1990).

Johnson and Johnson (1983) are well known for their contributions to cooperative learning. They emphasizes face-to-face interaction as students work in heterogeneous groups with individual accountability on tasks that require positive interdependence. Their guidelines for training students for cooperation are:

1. Make sure students see the need for the skill.
2. Make sure they understand what the skill is and when it is used.
3. Provide opportunities to practice and master the skill.
4. Give students feedback on their use of the skills and adequate time for skill development.
5. Make sure students practice the skill until it is internalized.
6. Have students process (evaluate) their use of the skills.

Teacher modeling of social skills and role playing are effective strategies for training students for cooperative learning. Johnson and Johnson (1983) identify four categories of skills to be developed for cooperative learning:

Forming Skills

- moving without noise
- staying within the group
- using quiet voices
- encouraging participation by all

Functioning Skills

- using names
- looking at the speaker
- being respectful of others
- directing the group's work
- expressing support
- asking for help or clarification
- offering to explain or clarify
- paraphrasing other people's work
- energizing the group
- describing feelings when appropriate

Formulating Skills

- summarizing out loud
- seeking accuracy by correcting and/or adding to summaries
- seeking elaboration
- seeking clever ways of remembering information
- demanding vocalization
- asking other members to plan out loud

Fermenting Skills

- criticizing ideas, not people
- integrating ideas
- asking for justification
- extending other students' answers
- asking in-depth questions
- generating further answers
- checking the group's work

Group processing is recommended as the final stage of a cooperative-learning group activity. Group processing refers to students reflecting on, analyzing, and evaluating how the group functioned. This can be done in at least two different ways. The group as a whole can evaluate its overall performance by systematically discussing issues such as:

Did everyone participate?
Did anyone dominate?
To what extent did group members stick to their assigned roles?
To what extent did students listen to each other carefully without interrupting each other?
To what extent were students courteous and respectful of each other?
How well did the group use its time?
To what extent did group members maintain eye contact while communicating?
What were the group's greatest strengths?
How could the group work better together next time?

Another way group processing can be conducted is to have a group discussion about the issues just described after students have individually assessed their own performance in the group. Once each group has evaluated its performance, the teacher can ask groups to share their results, list each group's strengths and weaknesses on the blackboard, and discuss as a class whether and what additional training in cooperative learning may be appropriate.

Structuring Groups

Although most proponents of cooperative learning in mathematics recommend student pairs or heterogeneous groups of three to five students, there is not total consensus on group size or structure. Some teachers begin with students working in pairs and gradually shift to three, then four, and finally five students working together. Many people recommend groups of four, because this size of group is large enough for generating ideas and discussing solutions of challenging problems, is small enough for all students to participate; can be conducted without a leader, and can be split up into pairs for occasional practice (Davidson, 1990). Groups may stay together for a single class period, a week, a school term, or the entire school year. Many proponents of cooperative learning recommend changing groups often enough that eventually all students in the class have the opportunity to work together. Some teachers allow students to choose their own groups; many assign students to groups to ensure heterogeneity. Assignment can be done on the basis of students' achievement records, through random grouping, or according to the regular seating chart. Substantial research supports the value of gender, racial, and ethnic diversity in cooperative-learning groups.

Deciding how to arrange groups in the classroom is another decision teachers have to make. Movable desks or tables and chairs are important but not absolutely necessary; with them, groups can be arranged so as to facilitate communication without disturbing other groups. Pairs can easily be used even in large lecture halls with no movable furniture or room to sit on the floor.

There are numerous ways to use cooperative learning for teaching problem solving. Some methods are relatively informal in structure but involve students working together cooperatively to achieve a problem-solving goal. Other methods are highly structured in terms of student roles, instructional techniques, materials, and assessment strategies. Well-structured cooperative-learning groups can increase the chances that all students will examine alternative solution strategies, will observe peers engaged in problem solving, and will formulate, analyze, and interpret problems and solutions. Positive interdependence and face-to-face interaction are guiding features behind some approaches to cooperative learning in mathematics. These are facilitated through assigning distinct roles to perform within groups (Johnson & Johnson, 1990).

Assigning Student Roles

Teachers sometimes have each group member fulfill a distinct role to ensure that all students participate in the process and to achieve a variety of goals. However, specific roles are not necessary. For example, the Small-Group Discovery Method of using cooperative learning in mathematics does not involve student roles. Some teachers prefer to assign students to their roles; other teachers prefer students to choose their own roles in the groups. Students should understand that their participation in the group is not limited to the roles they perform. Their roles are just one aspect of their participation in the group's problem-solving activities. Students need to see and be part of the "big picture" of problem solving. Many teachers have students change roles after one class period or after one group project. Regardless of which method is used, teachers should make sure that eventually all students get a chance to function in each of the group roles. There are several models of roles that can be use for problem-solving groups.

Model 1

- Problem restater: Paraphrases the problem and says what information is given and what must be found.
- Elaborator: Asks group members whether the problem is similar to others they have solved before.
- Strategy suggester/seeker: Makes suggestions about possible strategies that could be used to solve the problem, and/or asks groupmates for alternative strategies.
- ⑮ Approximator: Gets the group to estimate what the answer will be before they begin actually solving the problem.
- Reviewer/mistake manager: Has the group figure out how they can learn from whatever mistakes may have been made and, when the group is successful, has the group determine how their solution could be even better next time.
- Confidence builder: Encourages the group to keep going because they will succeed if they persist and work together effectively. (Johnson & Johnson, 1990)

Model 2

- Initiator: Gets the group started and keeps them on task.
- Idea person: Gives mathematical ideas to the group, such as how to solve the problem.
- Challenger or rebel: Does not passively accept the approaches selected and the answers but questions whether they are correct.
- ③ Synthesizer: Resolves differences by reconciling opposing views—the peacemaker.
- Ego builder: Builds pride in the group by praising its members. (Crabill, 1990)

Model 3

Jigsaw (Aronson et al., 1978) is a method of cooperative learning that involves assigning each student responsibility for solving a particular part of the problem, structuring material and activities so students become experts on their parts, and teaching their parts to the rest of the group. Problems need to be carefully selected so that students can solve their parts relatively independently.

Model 4

- Facilitator: Leads the group in the problem-solving process.
- Comprehension monitor/clarifier: Checks up on group members' understanding of the problem and how they are solving it; clarifies and/or asks for clarification as needed.

- Checker for accuracy and direction: Checks for mistakes, makes sure the approach they are using is leading in the right direction.
- Encourager: Praises group members and invites their contributions.
- Summarizer: Reviews progress the group has made and what still needs to be done.
- Recorder/reporter: Writes down the group's strategies, calculations, and answers, and presents this information to the class.

Model 5

- Accountant: Makes sure all group members perform mathematical operations.
- Architect: Ensures all group members contribute to the overall group product.
- Elaborator: Relates current work to prior mathematical problems. (Johnson & Johnson, 1989)

There at least two different types of group roles that can be assigned: management and instructional roles. Examples of management roles include facilitator and recorder; examples of instructional roles include chief investigator and adviser. Some controversy exists about the value and effectiveness of different types of roles. Management roles have been found to be especially important for younger children because they help prevent confusion, wasted time, and conflicts, while instructional roles sometimes create problems for both older and younger children (Good et al., 1989/1990).

Selecting or Preparing Materials

Although some teachers find they can adapt their existing mathematics curriculum for cooperative learning, others find that more appropriate materials are needed. There are several resources specially developed for implementing cooperative learning in mathematics. One especially good resource is *Get It Together: Mathematics Problems for Groups, Grades 4-12* (Erickson, 1989), developed by Project Equals at Lawrence Hall of Science, University of California, Berkeley.

The NCTM handbook on cooperative learning is a clear and concise guide with carefully constructed problems, clear instructional objectives, problem sheets, activities, and teaching notes (Artzt & Newman, 1990). Others include the Real Maths curricular materials (Terwel, 1990) in the Netherlands, the Finding Out/Descubrimeinto curriculum (Lotan & Benton, 1990) from the Stanford University, California, and *Developmental Mathematics* (Hackworth, Howland, & Alwin, 1989). The final chapter of the outstanding book by Davidson (1990), *Cooperative Learning in Mathematics*, identifies additional resource materials. *Designing Groupwork* (Cohen, 1984) contains excellent descriptions of several cooperative-learning methods and especially useful materials and procedures for training students for cooperation. There are other good sources, too numerous to mention, on using cooperative learning to teach mathematical problem solving.

Monitoring and Evaluating Groupwork

As a manager of learning, teachers using cooperative learning must monitor and evaluate groups as they solve problems. Monitoring involves checking up on students' performance of the assigned task. While cooperative learning is in progress, the teacher circulates from group to group observing and listening to students, and interpreting the results. It is best for the teacher to think about what will be looked for in the groups before the lesson begins. Johnson and Johnson (1975) describe monitoring as "the fun part" of cooperative learning for the teacher. Students can become enthusiastic problem solvers, and this is exciting to observe in progress. Teachers should check to see if students are on task, if they understand the assignment, if they are really working together cooperatively, if they are functioning in any roles that may have been assigned, and if there are any problems that should be dealt with. Intervention in groupwork by the teacher should be avoided unless absolutely necessary. If intervention is called for, it should come in the form of questions that guide students to resolve their own problems or that steer them in the right direction. Teachers should resist the temptation to tell students answers or how to solve problems. During this phase the teacher can provide feedback and praise students for good performance. While observing groups, the teacher may discover that students need additional information to help them solve a problem or additional training in working together cooperatively. Therefore, part of the monitoring process is to evaluate groupwork and to plan for follow-up instruction.

Teachers need to evaluate at least three aspects of group performance: (a) how students solved their problems (e.g., What strategies were considered? Which approaches were rejected? Why? Were there any careless mistakes? Did students verify their solutions?), (b) the solutions students generated, and (c) groups' processes (i.e. how the groups functioned as cooperative learners).

Some teachers offer rewards to groups for good performance. Competitions between groups are sometimes used as an incentive. Praising groups for good performance is usually an effective motivational tool.

An important decision teachers need to make is how they will evaluate individual student achievement of the specific objectives targeted by the lesson. Both individual and group accountability are important to emphasize. To ensure individual accountability, teachers can give tests (in-class or take-home) and quizzes or randomly ask group members to explain problem solutions. Additional measures include classwork (attendance, participation, cooperation), group projects, homework, self-evaluation, and peer evaluation (Davidson, 1990). When evaluating the group product, the teacher can give each student a grade for her or his individual contribution. Students should know the specific evaluation criteria that will be used to assess their performance.

The Potential Pitfalls of Cooperative Learning

Like other teaching methods, cooperative learning can be ineffective if it is not handled right. Not all groupwork is cooperative learning. Students can sit side by side in a group and do their work completely independently without cooperating. Potential student-oriented problems include a group of students becoming bored with each other, inadequate leadership within a group, students feeling abandoned by the teacher, difficult problems causing feelings of defeat, easy problems becoming boring, and students' need for a change of pace or more praise. Teacher-oriented problems include discomfort at not being the center of the classroom, inadequate explanation of the task, and mixed feedback about what students have learned (Crabill, 1990).

Although many students prefer working cooperatively to working independently, some students would rather work alone. Such students can inhibit effective group interaction. The teacher's role as observer and supervisor is important in this type of situation. Depending on the particular class and curriculum, teachers may decide to use cooperative learning as an option for students rather than as a requirement.

Another problem is that one or two students can do all the work solving problems while the others do not. To prevent this, individual accountability is essential. Groups must be structured to foster cooperation among students. Assigning roles, sharing materials, requiring a group product, and using group incentives can help structure effective cooperation.

Time can be a problem when implementing cooperative learning, and lessons sometimes end without summarizing what was learned and assessing the group process. One way of handling this problem is to assign roles of summarizer and leader of group processing. Otherwise the teacher can lead the summarizing and group processing at the end of each lesson (or set of lessons).

Cooperative learning in mathematics can lead to incoherent presentation and interruptions while working so that students need to spend more time reviewing and practicing. If initial training of students to work cooperatively is not adequate for some students or groups, follow-up training may be needed. Sometimes problems arise if teachers set only academic goals rather than specifying both academic and social goals in advance (Good et al., 1989/1990).

Conclusion

Cooperative learning has many advantages for teaching problem solving. It improves math achievement and higher-level mathematical thinking. Working in pairs or small groups is highly motivating for most students and improves their attitudes about themselves as learners and problem solvers. Improved interpersonal relations among students of different ethnic/cultural backgrounds is another benefit of cooperative learning.

With its 20-plus years of success as a technique for developing mathematical problem-solving skills, cooperative learning is not just a fad. The benefits of using cooperative learning as a technique for problem solving have been well documented, and they far outweigh the disadvantages. Teachers can develop personal action plans to design cooperative learning lessons that meet the needs of their specific students and curriculum. However a teacher might adapt cooperative learning to teach mathematics, it is likely to make problem solving more lively and fun for both teachers and students.

Note

1. I am grateful to Elaine Giardino for her contributions to the development of this example.

References

Aronson, E., Blaney, N. Stephan, C., Sikes, J., & Snapp, M. (1978). *The jigsaw classroom.* Beverly Hills, CA: Sage.

Artzt, A., & Armour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of group problem solving in mathematics. *Cognition and Instruction, 9*(2), 137-175.

Artzt, A., & Newman, C. (1992). *How to use cooperative learning in the mathematics class.* Reston, VA: National Council of Teachers of Mathematics.

Bransford, J., & Stein, B. (1984). *The IDEAL problem solver.* New York: Freeman.

Burns, M. (1990). The math solution: Using groups of four. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Chemeketa Community College. (1986). *The 6PQ method of discovery learning.* Salem, OR: Author.

Cohen, E. (1984). *Designing groupwork: Strategies for heterogeneous classrooms.* New York: Teacher's College Press.

Crabill, C. (1990). Small-group learning in the secondary mathematics classroom. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Davidson, N. (1990). The small-group discovery method in secondary and college level mathematics. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Erickson, T. (1989). *Get it together: Mathematics problems for groups grades 4-12.* Tucson, AZ: Zephyr.

Fantuzzo, J., Alperin King, J., Rio Heller, L. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment. *Journal of Educational Psychology, 84*(3), 331-339.

Frankenstein, M. (1986, Spring/Summer). Ideas for teaching a non-rote college arithmetic course. *Mathematics in College,* pp. 22-31.

Good, T., Reys, B., Grouws, D., & Mulryan, C. (1989/1990). Using work-groups in mathematics instruction. *Educational Leadership, 47*(4), 56-62.

Gourgey, A. (1992). Tutoring developmental mathematics: Overcoming anxiety and fostering independent learning. *Journal of Developmental Education, 15*(3), 10-14.

Hackworth, R., Howland, J., & Alwin, R. (1989). *Developmental mathematics.* Clearwater, FL: H & H Publishing.

Hartman, H. (1992). *Intelligent tutoring.* Clearwater, FL: H & H Publishing.

Heiman, M., Narode, R., Slomianko, J., & Lochhead, J. (1987a). *Teaching thinking skills: Mathematics.* Washington, DC: National Education Association.

Heiman, M., Narode, R., Slomianko, J., & Lochhead, J. (1987b). *Teaching thinking skills: Science.* Washington, DC: National Education Association.

Johnson, D., & Johnson, R. (1975). *Learning together and alone.* Englewood Cliffs, NJ: Prentice Hall.

Johnson, D., & Johnson, R. (1983). *Overview of cooperative learning.* Minneapolis: University of Minnesota.

Johnson, D., & Johnson, R. (1990). Using cooperative learning in mathematics. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Lotan, R., & Benton, J. (1990). Finding out about complex instruction: Teaching math and science in heterogeneous classrooms. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Pressley, M., Wood, E., Woloshuyn, V., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*(1), 91-109.

Rosenshine, B., & Meister, C. (1992, April). The use of scaffolds for teaching higher-level cognitive strategies. *Educational Leadership,* pp. 26-33.

Schoenfeld, A. (1989). Teaching mathematical thinking and problem solving. In *Toward a thinking curriculum: Current cognitive research.* Alexandria, VA: Association for Supervision and Curriculum Development Yearbook.

Slavin, R. (1990). *Cooperative learning research, theory and practice.* Englewood Cliffs, NJ: Prentice Hall.

Terwel, J. (1990). Real maths in cooperative groups in secondary education. In N. Davidson (Ed.), *Cooperative learning in mathematics.* Tucson, AZ: Zephyr.

Webb, N. (1985). Student interaction and learning in small groups: A research summary. In R. E. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, C. Webb, & R. Schmuch (Eds.), *Learning to cooperate, cooperating to learn.* New York: Plenum.

Whimbey, A., & Lochhead, J. (1982). *Problem solving and comprehension.* Philadelphia: Franklin Institute Press.

Zhu, X., & Simon, H. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*(3), 137-166.

Recommended Reading

Baron, J., & Sternberg. (Eds.). (1987). *Teaching thinking skills.* New York: Freeman.

Barell, J. (1991). *Teaching for thoughtfulness.* White Plains, NY: Longman.

Chipman, S., Segal, J., & Glaser, R. (Eds.). (1985). *Thinking and learning skills.* Hillsdale, NJ: Lawrence Erlbaum.

Wertheimer, M. (1959). *Productive thinking* (Enlarged ed.). New York: Harper & Row.

20

Problem Solving and the Mathematically Gifted Student

A Psychological Perspective

BRIGITTE A. ROLLETT

In spite of the fact that the development of higher mathematical ability has now been studied intensively for almost a century, there is still a broad range of intriguing questions that have yet to be investigated. Foremost amongst these are the causes for the highly gifted individuals' exceptionally efficient ways of processing information. Of special importance is the neurophysiological basis of cognitive and personality development and the interaction of the related factors. Some insight into problem solving can be obtained from an understanding of these phenomena.

Defining Giftedness

The definition of *giftedness* is still somewhat contentious in the literature. However, clear-cut diagnoses of giftedness, which might form the prerequisite for a selective early stimulation of exceptionally competent students, are only possible once we can proceed from exact definitions of excellency. Sternberg and Davidson (1986) identified 17 concepts of giftedness, connected with one another in specific ways but differing with regard to fundamental features. In their own approach they accorded "insight," a particularly efficient form of processing information and solving problems first described by Wertheimer (1945), central importance in their definition of intellectual giftedness.

Important aspects they identified are:

1. *Selective encoding:* differentiating important aspects of a problem from unimportant ones
2. *Selective combination:* efficient fusing of isolated facts, structuring them in such a way as to be helpful in solving the problem at hand
3. *Selective comparison:* comparing new insights with previous knowledge about the problem and developing higher-order structures

Howard Gardner (1983) formulated the theory of "multiple intelligences," which proceeds from the premise that the abilities to solve a problem competently in a specific field (e.g., mathematics, music, art, different areas of science, etc.) develop relatively independently of one another and lead to qualitatively different forms of intelligence. Even in 18-month-olds, for example, different gifts can be observed. The technically gifted "patterners" prefer playing with building blocks and other construction toys, thus training their technical faculties, while the linguistically gifted "dramatists" like playing with dolls and toy animals and in this way develop their linguistic and communication talents (Shotwell, Wolf, & Gardner, 1979).

For this reason we have turned to a dynamic concept of giftedness today. Giftedness can only unfold if there is a beneficial interaction with the environment. This is in keeping with the new theories of development, which, when explaining human behavior, no longer proceed exclusively from a theory of heredity or exclusively from one of environment but assume an interaction between both components. Inherent potential and a conducive environment must combine in the development of supreme achievement.

Early Indications of Exceptional Ability

Children who are later identified as gifted often show interest in their environment at a very early age. Even as newborn infants, they are "efficient orienters," as Stapf and Stapf (1988) established in an investigation of 51 gifted children. The orienting reaction is an indicator of the child's readiness to learn from his or her environment. This active information-seeking behavior is optimal in furthering the development of the functional units of the brain. One of the basic observations made when comparing efficient learners with their less able peers is that the former seem to be better able to identify relevant information from unimportant data and to concentrate on processing it.

We therefore want to define giftedness as the ability expertly to identify, construct, process, and handle relevant information in one or more culturally defined fields of action. Renzulli's classic model of giftedness further elaborates this by drawing attention to the role creativity (the ability to produce relevant new ideas) and task commitment (willingness to work toward high accomplishment) play in attaining superior results.

It is not surprising that novelty preference is a better predictor of cognitive development than the parents' educational level, as Rose and Wallace (1985) showed in a longitudinal study of the cognitive functioning of preschool children.

Highly gifted children need less time to react to and process information. As most functions of the cerebral cortex only develop after birth in response to stimulation by the environment, this interest in new information brings about a significant leap forward in the child's development, one which can, of course, be increased if the child's persons of reference know how to play with him or her in a stimulating way and to encourage it. It is typical of mathematical giftedness to appear, together with high verbal proficiency, at a very early age (3 to 4 years). Early mastery of the symbolic nature of language seems to be a milestone in intellectual giftedness. Another is the early specialization of the functions of the brain.

Degree of Lateralization of Cerebral Functions and Giftedness

The general interpretive area, or "Wernicke's area," in the brain plays the most important role in the development of language, including dealing with numbers and the handling of meaningful information in general. Consciousness and self-awareness are mediated through this area. Thus, it represents the core of the personality and its development.

As one hemisphere—usually the left one—is principally used when processing new information, it gains dominance over the other hemisphere, becoming more and more expert in dealing with new learning experiences and their results. This process of lateralization leads to an efficient division of labor between the two hemispheres, as Nobel laureate Sperry (1974) showed: With most people, the left hemisphere becomes devoted to the processing of language and numbers and, later, the development of the writing and reading centers, while the right hemisphere is dedicated to spatial orientation and construction, the processing of pictorial and color information, music, and the understanding and interpreting of higher-order emotions. From this description it is clear that both hemispheres have to work together when solving mathematical problems, as Levin and Spiers (1985) in their study of acalculia showed. While the left hemisphere handles the numerical or digital information to be computed, the right one processes the schematic, pictorial, or analog part of a problem.

Gifted individuals not only show a higher degree of lateralization and, in consequence, specialization of the two hemispheres, but are also more proficient in integrating the messages of the "two brains."

Creativity and Giftedness

Creativity, the ability to make unusual associations of ideas and find uncommon solutions, presents a special problem (see Urban, 1992). It is central to extraordinary achievement in many fields and to success in later life but is rarely stimulated by parents and educators. These people have to be creative themselves in order to be able to recognize creative attainments as such. Parents and educators also need to react in a suitable way, that is, by helping to develop the idea in a playful manner or simply by acknowledging its correctness, as Gauss's teacher did when the young Gauss presented him with the ingenious solution for the problem of adding all the numbers from 1 to 100 by multiplying 101 by 50, thus intuitively deducing the principle of the arithmetical progression.

At the age of about 4, most children make a "creativity jump." At this age, language, especially, develops; but also the imagination begins, and with it, joy in new, surprising associations of thoughts and words, new word creations, and creative, imaginative games, including games with numbers. Depending on whether this development is fostered by the environment or dismissed as "nonsense," it can grow into a stable feature of character or merely atrophy.

Effects of Accelerating the Gifted

Recently, there has been much discussion whether or not to accelerate the gifted. Renzulli and McGreevy (1986) investigated 23 pairs of twins, of whom only one twin in each case had been diagnosed as being gifted and admitted to a special training program. It was seen that the other twin later indeed displayed a poorer academic performance but was much more creative and independent in thought.

In the course of their study of mathematically precocious youths, Swiatek and Benbow (1991) compared accelerated gifted mathematics students with others who were not accelerated, following their academic and psychosocial development during a 10-year period. As was to be expected, the accelerated group finished college 1 year earlier. Both groups showed outstanding academic and personal satisfaction at age 23. The authors concluded that the accelerates had only a slight advantage over the unaccelerated. However, an analysis of their data in terms of effect sizes shows a clear shift of the accelerates toward academic preferences, while the unaccelerated group had more diverse interests and developed higher self-esteem, presumably because they were not continually exposed to academic peer pressure. Whether one advocates acceleration of the gifted or simply provides them with an enriched curriculum depends on the educational goals one prefers. One such form of enrichment would be the integrated inclusion of a problem-solving strand in the syllabus.

Detrimental Effects of One-Sided Stimulation

A one-sided stimulation can be detrimental when the aim is training basic skills in a specific field, but the opportunity is not taken to introduce the child to the more demanding areas of the subject. This can happen very easily with mathematically gifted children.

Unusual mathematical abilities can be discerned very early in life in a pronounced interest in numbers and their relationships. Parents and educators, who probably are not expert in mathematics, are pleased with the child's arithmetical stunts but are not in a position to familiarize the child, in a playful manner, with concepts in mathematics. This results in mathematics coming to a standstill at the level of the "mental calculator," of the acrobat in mental arithmetic. Hope (1987) cites a case in point. At the time of the investigation 13 years earlier as a small child, Charlene had already discovered that it was fun "to play with numbers." She then developed exceptional abilities in mental arithmetic. She was able to calculate 87 × 23 correctly at lightning speed. On being asked, she stated that she had restructured the problem as follows. She had immediately realized that 87 corresponds to the product of 3 × 29 and had made the following mental rearrangement: $87 \times 23 = (29 \times 3) \times 23 = 29 \times (3 \times 23) = 29 \times 69 = 69 \times (30 - 1)$. So all that remained for her to do was the simple calculation $69 \times 30 - 69$, and she came to the correct solution, 2001.

It is clear that a lengthy preoccupation with numbers and their qualities and the extensive, easily addressable internal knowledge base deriving from it are necessary to sense such a calculation as a "relief." However, Charlene had received no further encouragement and had thus come to a standstill at the level of perfect mental arithmetic, something she had taught herself. She had found no access to higher mathematics. This case shows dramatically that giftedness alone is not enough; finely attuned stimulation of the individual case is indispensable.

The case of the Indian boy Balamurati Krishna Ambati represents an example demonstrating the opposite. At the age of 4 he possessed extraordinary mathematical abilities, received proper support, and at the age of 10 achieved the impressive score of 750 on the SAT, ensuring him a place in college (Papalia & Wendkos Olds, 1992, p. 269).

Encouraging the Gifted

Renzulli (1980) introduced another important aspect of encouragement into the discussion. In his model, a high degree of commitment—apart from general above-average abilities and a high level of creativity—is necessary to realize the gift. This means that the creation of a positive achievement motivation, one characterized by pleasure in success, must be a central concern in any stimulation of giftedness. Mönks and Boxtel (1985) added to this model by including the agents of stimulation (the family, school, the self). Supreme achievement is possible only when parents and teachers manage to make the activity in question pleasant for the child so that a stable intrinsic motivation can develop.

The younger the child, the easier it is to spoil for them a potentially stimulating activity, if the learning atmosphere is unpleasant or they are overtaxed. As W. W. Sawyer (1943) wrote in his truely enjoyable *Mathematician's Delight*, "To master anything—from football to relativity—requires effort. But it does *not* require *unpleasant* effort, drudgery" (p. 9).

Thus, highly gifted children need not only a stimulating but also a nurturing and encouraging social environment to fulfill their potential and develop into emotionally and psychologically well-adjusted adults.

Developing Metacognition

Metacognition has been a buzzword now for some time. As a general competency it begins to appear at about 8 years of age, accelerating its development at about 11, when, according to Piaget, formal logic and abstract thinking develop. Gifted children show precursors of metacognitive thinking at a very early age.

Shore and Dover (1987) have been able to demonstrate that there are characteristic differences between gifted and normal children with regard to the use of metacognitive strategies. Metacognition refers to thinking about thinking, that is, rational deliberations about the best strategic employment of one's own competences and strategies to solve problems. Gifted children manage to guide and thus optimize their handling of problems through metacognition earlier and more systematically.

The way parents and teachers treat the child when solving problems together has proved to be of special significance to the development of cognitive competences. Parental interaction with the child when dealing with problems is the scaffolding for the development of intelligent strategies to solve problems. In the course of dealing with problems in conjunction with parents, the child experiences more and more success in constructing his or her own solutions, thus developing his or her intellectual and metacognitive competences.

The difference in information processing among gifted, normal, and slow students is well illustrated by a study conducted by Lehtinen (1992, p. 127), who asked students to summarize a mathematical lesson dealing with changing fractions into decimals, writing down what was most important.

A gifted girl wrote: "During this lesson we were taught many important things about turning a fraction into decimals, where the denominator is 10, 100, or 1000, e.g., $\frac{2}{10} = 0.2$ and $\frac{21}{100} = .21$, and the multiples of a fraction, whose denominator is not 10, 100, or 1000, and their transformation into a decimal, e.g., $\frac{4}{25} = \frac{16}{100} = 0.16$; $\frac{2}{5} = \frac{4}{10} = 0.4$; $3\frac{2}{5} = 3\frac{4}{10} = 3.4$."

This girl knew how to identify the relevant core information of the lesson and construct a well-ordered information module to be remembered, thus augmenting her own internal knowledge base, by illustrating the procedures by a few well-chosen examples.

On the other hand, one average student, wrote: "During this lesson that [which was important] was taking multiples, e.g., $\frac{5}{2} = \frac{25}{10} = 2.5$. This transforming into a decimal was the most important, e.g., $\frac{50}{100} = 0.50$, and telling about it was the most important."

While this description does not include any false information, it is certainly not complete and it lacks structure.

The third example was produced by a willing but very slow learner: "This lesson. During this lesson one was supposed to listen and ask questions. $\frac{10}{2} = 0.2$. Without instruction I would not have known this." (On the blackboard, the teacher had written "$\frac{2}{10} = 0.2$.")

During the lesson, this student had willingly played the "question-and-answer game," but he was unable to correctly reconstruct the content of the lesson for himself. He would have needed far more individual encouragement and information feedback to arrive at a workable knowledge base enabling him to keep pace with his peers.

Characteristics of Gifted Math Students

There is much to be learned from teaching a normal mathematics class by studying the way gifted students approach a mathematical problem. Successful mathematics students have several characteristics in common. At an early age, they discover that mathematics is a special language allowing them to detect, describe, and draw inferences from regularities, or patterns. This opens up a fascinating world of problems and solutions, and they are usually eager to explore its possibilities.

They show a high degree of intellectual curiosity and willingness to work on a challenging problem for its own sake, not because the teacher demands it or they want to get good grades. In their thinking, they proceed in a methodical, metacognitively governed way. They know how to ask themselves pertinent questions to get a clear view of the problem. Polya (1988, p. 2) lists some of the most important ones for understanding a problem: "What is the unknown? What are the data? What is the condition?" This provides a good guide to follow when teaching mathematics.

When working on a problem, gifted students resort to a procedure closely resembling Polya's four phases:

1. Strive to *understand the problem* and find out what is required
2. Decide how the information has to be reorganized to get an idea of the solution, and *devise at a plan*, try it out, and revise it if necessary.

3. *Carry it out.*
4. Most important, *look back* and check the results not only for corrections, but also for determining possible further use (Polya, 1988, pp. xvi-xvii).

Using videotapes of their problem-solving processes, Eva Schneider (1992) compared pairs of good math students with pairs of poor math students. Using Schoenfeld's categories of mathematical problem-solving behavior—resources, heuristics, control, and belief systems—she found an important difference: Pairs of sucessful math students used significantly more control strategies.

In an early study of the learning styles of seventh and eighth graders, we found that the most successful students typically did *not* just study for a given exam only to forget all about the whole thing after they had passed it; rather, they endeavored to store the new insights in a well-structured form into their internal "knowledge base." This not only saved them a lot of time and effort when a similar problem had to be dealt with, but it also gave them a head start when working on a new problem, because they could use the general strategies that had already served them well. Learning for future use and striving to build up a well-structured and easily addressable knowledge base seem to be important characteristics of gifted students.

Some Practical Examples and Their Implications for Teaching Mathematics

Gifted students know how to use intelligent visualization of problems as a metacognitive strategy; they are able to translate verbal information into pictorial information, and vice versa, to get a better grasp of the problem.

Herwig Schatzl (personal communication, 1993), a mathematics teacher at an Austrian high school, relates the following example:

> I particulary remember from my teaching experience, a case regarding finding the sum of the series. Specifically, finding the sum of an infinite geometric progression.
>
> $$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots$$
>
> It is understood that this sum is the limit of the series of partial sums. We then get the formula for the limit $S_\infty = \lim S_n = a/(1-r)$, where a is the first term and r the common ratio.
>
> It is not always easy for students to see the connections and recognize the fact that the sum of this infinite series can be set equal to a particular value. In the above case, the result is $1+\frac{1}{2}+\frac{1}{4}+\ldots=2$.
>
> One 16-year-old student's argument at the time with the help of a number line

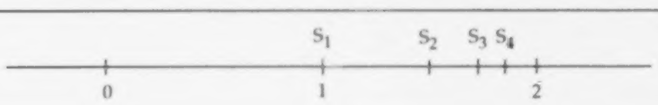


Figure 20.1

The following partial sum, S_{n+1} can always be obtained by taking half the difference of S_n and 2. In this way, the limit value of the sequence of partial sums clearly becomes 2. In this way, the general formula $S_\infty = a/(1-r)$ can be tested, and the notion of a limit, in particular the process of infinite convergence, was made particularly clear and apparent to the students.

It is important to note that this was a spontaneous creative insight into the problem by this student and not a prepared part of the teacher's lesson.

Here is another instance of the way gifted students reformulate a problem structure. Eva Turner gave the following problem to her 10th-grade mathematics class:

> An equilateral triangle with side length $a_1 = 8$ cm is inscribed in a circle. The inscribed circle of this equilateral triangle becomes the circumcircle of equilateral triangle a_2, and so on. The problem is to find the sum of the triangles' perimeters.

The following is an outline of the teacher's solutions:

Note: The largest triangle's perimeter is U_1, its side length is a_1, its in radius is r_1, and its altitude is h_1.

$$u_1 = 3a_1 = 24$$

$$h_1 = \frac{a_1}{2}\sqrt{3}$$

$$r_1 = \frac{1}{3}h_1 = \frac{a_1}{2}\cdot\frac{\sqrt{3}}{3}$$

$$a_2 : a_1 = r_1 : (h_1 - r_1)$$

$$a_2 = \frac{a_1 r_1}{h_1 - r_1}$$

$$u_2 = 3a_2 = 12$$

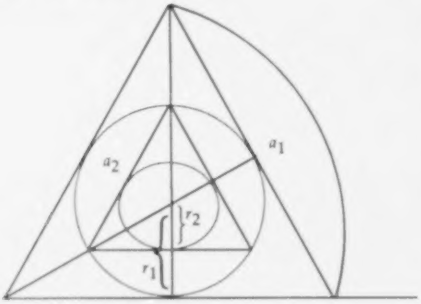


Figure 20.2

$\frac{u_2}{u_1} = q;\ q = \frac{1}{2}$. The common ratio between consecutive perimeters is $\frac{1}{2}$; therefore the required sum is s.

$$s = 24\cdot\frac{1}{\frac{1}{2}} = 48$$

In this case, the student realized that he could use the tangent function he had previously learned to arrive at a more elegant solution:

$$a_1 = 8 \text{ cm} \Rightarrow u_1 = 24 \text{ cm}$$

$$r_1 = \frac{a_1}{2} \tan 30^\circ \qquad r_1 = \frac{2}{3} \cdot h_2 = \frac{2}{3} \cdot \frac{a_2}{2}\sqrt{3} = \frac{a_2}{3}\sqrt{3}$$

$$a_2 = \frac{3r_1}{\sqrt{3}};\ a_2 = 4\text{cm} \Rightarrow u_2 = 12 \text{ cm}$$

As the teacher explained: "Normally, students at age 15 have not learned the tangent function and therefore do not have it at their disposal. However, this year I chose to begin with trigonometry and then did the required work in the proper order" (Turner, personal communication, 1993).

It is typical for gifted students to integrate new information into their existing knowledge base, forming a new gestalt that makes it possible to see how different modules of information relate to each other and can be used to advantage. Average students tend to "compartmentalize" what they know. It is therefore a good teaching strategy to work against this tendency, to let students deduce relations between different modules of information, and to show them how they can make use of them.

Intelligent students often try to solve problems according to a "least-effort principle," finding shortcuts and producing elegant solutions. Another example Eva Turner relates begins with this problem:

> The sum of three numbers in an arithmetic progression is 24. The sum of the squares is 642. Find the three numbers.

$$a_1 + (a_1 + d) + (a_1 + 2d) = 24 \qquad a_1^2 + (a_1 + d)^2 + (a_1 + 2d)^2 = 642$$

$$a_1 = 8 - d \qquad 3a_1^2 + 6a_1d + 5d^2 = 642$$

By substitution of a_1 in the last equation, the value for d is obtained: $d = \pm 15$. yielding the arithmetic progression –7, 3, 23 or 23, 8, –7.

A solution produced by a talented student began with the three numbers selected differently:

$$a_2 - d,\ a_2,\ a_2 + d$$

so that

$$(a_2 - d) + a_2 + (a_2 + d) = 24,$$

and

$$3a_2 = 24$$

$$a_2 = 8$$

$$a_2^2 - 2a_2d + d^2 + a_2^2 + a_2^2 + 2a_2d + d^2 = 642$$

$$2d^2 = 450$$

$$d = \pm 15,$$

a much more elegant solution.

To arrive at such an insightful solution, it is necessary to "see" the structure of a problem and realize its potentials for reorganization. How is this accomplished? A well-structured knowledge base is, of course, essential, but it is also important to give oneself the time to ponder the problem before starting on it in the standard way. Average students tend to be so engrossed with the "standard solution" that they start on it right away once they have hit on it, without giving themselves time to reflect on alternate, more elegant solutions. The teacher can counteract this by allocating enough time for her class to think (not forgetting to provide enrichment for gifted students who would be bored by having to wait for the rest of the class to come to terms with a given problem) and encouraging them to try to find new, intelligent solutions.

This does of course not explain the lightning speed at which gifted students sometimes come up with a brilliant idea. What happens in such a case is that gifted students have managed to automate much of their thinking, so that it suffices for them to thoroughly understand a given problem and to send a "search order" into the processing units of the brain to come up with a solution that meets the criterion they have decided on. Only when asked to describe the procedures do they take the trouble to recount every step of their reasoning (as a demonstration, see Charlene's explanation of her mental calculation, above). According to Galperin's action theory, this automatization of internal thinking processes is a normal occurrence, resembling the change from direct machine programming of computers to using special program packages: The expert thinker no longer has to spell out every step but automatically processes big chunks of information by addressing them by an appropriate thought, comparing the results, and refuting or using them as may be the case with hardly any conscious cogitation.

The story of McKay's theorem is an example for such a seemingly effortless solution found by an eighth grader cited by Bruckheimer and Markovits (1986):

> Scherzer (1973) describes how he was discussing the "betweenness" property of rational numbers and how to find a fraction between any two given fractions. He had worked through a number of examples to

> find the fraction midway between two fractions when, to his consternation, McKay, one of the students, interrupted, "Sir, you don't have to go to all that trouble to find a fraction between two fractions, all you have to do is add the tops and bottoms."
>
> Of course, McKay was right:
>
> If $0 < \frac{a}{b} < \frac{c}{d}$, then $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$, which is McKay's theorem. (pp. 294-295)

The ease with which gifted students can find new and labor-saving procedures may, however, sometimes turn against them, when they shirk the trouble of actually working through the problem. The following story by high school mathematics teacher Richard Henner elucidates this problem:

> After the nuclear reactor catastrophy at Chernobyl on 26 April 1986, there were lots of radioactive isotopes ^{134}Cs (whose half-life is 2.3 years), spread by the winds throughout Europe. Develop an equation to show the radioactive breakup (for an initial mass of m_0). Draw the function, using $m_0 = 10$ cm.
>
> Calculate and check, with the help of your graph, what percentage of the isotopes since the accident occurred (about 6.9 years) has broken down. How long will it take until 99% of the beginning mass has broken down?

After being given the problem to develop the equation and use appropriate methods to answer the various questions, all the students very diligently completed the work. All students saved the most time-consuming work, drawing the equation, for the end—except for Johanna, a very gifted and ambitious, though not always very diligent, girl. She began to solve the problem by using the graph, which in large measure she did incorrectly. Because she did not have the equation of the function at her disposal, she had only the half-life values to use in graphing the curve. In the process, she noticed something I had overlooked in making the assignment—coincidentally a triple half-life had passed since the date of the accident. Therefore, a good part of the assignment was, for her, simply mental arithmatic. One eighth of a substance still exists, and therefore 3/8 still needs to be disposed of.

In this case, an "effort avoidance motive" (Rollett, 1987) led the student to find this elegant solution. This motive, however, can prove to be a double-edged sword, as her teacher comments:

> Incidentally, Johanna and similar students are amongst my biggest problem cases. I have often experienced that children at these grade levels, at age 14 and 15, because of the praise they receive for their insightful ideas, seem to stop practicing and doing homework, and simply wait for the ideas to occur. Creative students fail at substantive investigation in mathematics courses preparing for the Olympiads, as they are satisfied with understanding the problem and recognizing a clever way to a solution as opposed to going through an exact solution which may be too time-consuming.

As the story of Fermat's (1601-1665) famous "last theorem" shows, this reluctance of brilliant thinkers to work out their solutions in detail may have delayed discovery of the proof about 350 years.

The conclusion can only be that giftedness per se does not guarantee success, but needs skillful mentoring, including praise not only for brilliant ideas but also for diligently realizing them. Also crucial is a supporting environment that ensures that mathematical problem solving remains a rewarding and intellectually stimulating experience.

Since enhancing individuals' mathematical knowledge base plays such an important role in ensuring progress, students should be encouraged to learn not simply toward the goal of passing exams, but to strive to augment their own repertoire of strategies and insights and, to use Schoenfeld's terms, resources and heuristics—in short, learn to learn for future use.

The teacher should not insist on rote learning and standard procedures but allow the students to find new solutions, if this is feasible, and teach them how to control what they have achieved, actively using metacognitive techniques.

Another important characteristic of expert mathematicians that may lead to a useful strategy for efficient math instruction is their not being afraid of making mistakes. As Sawyer (1943) admonishes his readers, "The important thing is to learn how to strike out for yourself. Any mistakes you make can be corrected later. If you start by trying to be perfect, you will get nowhere. The road to perfection is by way of making mistakes" (p. 24). Only those who are willing to risk trying a seemingly useful procedure, but are aware of the possible need to change their approach, can become active, self-reliant explorers in the fascinating field of mathematics.

References

Bruckheimer, M., & Markovits, Z. (1986). Several incarnations of McKay. *Mathematics Teacher, 4,* 294-295.

Gardner, H. (1983). *Frames of mind.* New York: Basic Books.

Hope, J. A. (1987). A case study of a highly skilled mental calsulator. *Journal for Research in Mathematic Education, 18,* 331-342.

Lehtinen, E. (1992). Lern- und Bewältigungsstrategien im Unterricht. In H. Mandl & H. F. Friedrich (Eds.), *Lern- und Denkstrategien: Analyse und Intervention* (pp. 125-149). Göttingen: Hogrefe.

Levin, H., & Spiers, P. (1985). Acalculia. In K. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (pp. 97-114). New York: Oxford University Press.

Mönks, F., & Boxtel, H. W. (1985). Gifted adolescents: A developmental perspective. In J. Freeman (Ed.), *The psychology of gifted children* (pp. 275-295). New York: John Wiley.

Papalia, D. E., & Wendkos Olds, S. (1992). *Human development*. New York: McGraw-Hill.

Polya, G. (1988). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press.

Renzulli, J. S. (1980). What we don't know about programming for the gifted and talented. *PhiDelta Kappan, 61*, 601-602.

Renzulli, J. S., & McGreevy, A. M. (1986). Twins included and not included in special programs for the gifted. *Roeper Review, 9*(2), 120-127.

Rollett, B. (1987). Diagnosis and intervention in education and therapy. In K. Hurrelmann, F.-X. Kaufmann, & F. Lösel (Eds.), *Social intervention: Potential and constraints* (pp. 241-251). New York: de Gruyter.

Rose, S. A., & Wallace, I. F. (1985). Visual recognition memory: A predictor of later cognitive functioning in preterms. *Child Development, 56*, 843-852.

Sawyer, W. W. (1943). *Mathematician's delight*. Baltimore, MD: Penguin.

Schneider, E. (1992). *Analyse mathematischer Problemlösungsprozesse*. Unpublished master's thesis, University of Vienna, Institute of Psychology.

Shore, B., & Dover, A. (1987). Metacognition, intelligence and giftedness. *Gifted Child Quarterly, 31*(1), 37-39.

Shotwell, J. M., Wolf, D., & Gardner, H. (1979). Exploring early symbolization: Styles of achievement. In B. Sutton-Smith (Ed.), *Play and learning* (pp. 127-156). New York: Gardner.

Sperry, R. W. (1974). Lateral specialization in surgically separated hemispheres. In F. O. Schmitt & F. G. Worden (Hrsg.), *The neurosciences*. Cambridge: MIT Press.

Stapf, A., & Stapf, K. H. (1988). Kindliche Hochbegabung in entwicklungspsychologischer Sicht. *Psychologie in Erziehung und Unterricht, 35*, 1-17.

Sternberg, R. J., & Davidson, J. E. (1986). Conceptions of giftedness: A map of the terrain. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 3-18). New York: Cambridge University Press.

Swiatek, M. A., & Benbow, C. P. (1991). Ten-year longitudinal follow-up of ability-matched accelerated and unaccelerated gifted students. *Journal of Educational Psychology, 4*, 528-538.

Urban, K. (1992, November). *Recent trends in creativity research and theory*. Paper presented at the 3rd ECHA Conference, Munich, Germany.

Wertheimer, M. (1945). *Productive thinking*. New York: Harper & Brothers.

Sources for Problems

Abraham, R. M. (1932). *Winter nights entertainments*. London: Constable. (Reprinted as: *Easy-to-do entertainments and diversions with coins, cards, string, paper and matches*. New York: Dover, 1961.)

ApSimon, H. (1984). *Mathematical byways*. New York: Oxford University Press.

ApSimon, H. (1990). *More mathematical byways in ayling, beeling and ceiling*. New York: Oxford University Press.

Ainley, S. (1977). *Mathematical puzzles*. London: Bell.

Alcuin (attrib.). *Propositiones Alcuini doctoris Caroli Magni Imperatoris ad acuendos juvenes*. (Translated and annotated by John Hadley and David Singmaster as: Problems to sharpen the young. *Math. Gaz., 76*(No. 475) (1992), 102-126.)

Alexanderson, G. L., Klosinski, L. F., & Larson, L. C. (1985). *The William Lowell Putnam Mathematical Competition—Problems and solutions: 1965-1984*. Washington, DC: Mathematical Association of America.

Allen, L. (1991). *Brainsharpeners*. London: New English Library (Hodder & Stoughton).

Aref, M. N., & Wernick, W. (1986). *Problems and solutions in Euclidian geometry*. New York: Dover.

Artino, R. A., Galione, A. N., & Shell, N. (1982). *The contest problem book IV*. Washington, DC: Mathematical Association of America.

Barbeau, E., Klamkin, M., & Moser, W. (1976). *1001 problems in high school mathematics*. Montreal: Canadian Mathematical Congress. (Also reprinted in 1978, 1980, 1985)

Barr, S. (1965). *A miscellany of puzzles*. New York: Crowell.

Barr, S. (1969). *Second miscellany of puzzles*. New York: Macmillan. (Reissued as: *Mathematical brain benders*. New York: Dover, 1982.)

Barry, D. T., & Lux, J. R. (1984). *The Philips Academy Prize Examination in mathematics*. Palo Alto, CA: Dale Seymour Publications.

Bates, N. B., & Smith, S. M. (1980). *101 puzzle problems.* Concord, MA: Bates Publishing.

Berloquin, P. (1976). *100 geometric games.* New York: Scribner's. (Retitled as: *Geometric games.* London: Unwin, 1980).

Berloquin, P. (1976). *100 numerical games.* New York: Scribner's.

Berloquin, P. (1977). *100 games of logic.* New York: Scribner's. (Retitled as: *Games of logic.* London: Unwin, 1980.)

Berloquin, P. (1985). *The garden of the sphinx.* New York: Scribner's.

Birtwistle, C. (1971). *Mathematical puzzles and perplexities.* London: Allen & Unwin.

Brandes, L. G. (1975). *The math wizard* (Rev. ed.). Portland, ME: J. Weston Walch.

Bridgman, G. (1981). *Lake Wobegon math problems* (Rev. and enlarged ed.). Minneapolis: Author.

Brousseau, A. (1972). *Saint Mary's college mathematics contest problems.* Palo Alto, CA: Creative Publications.

Bryant, S. J., Graham, G. E., & Wiley, K. G. (1965). *Nonroutine problems in algebra, geometry, and trigonometry.* New York: McGraw-Hill.

Bryant, V., & Postill, R. (1980). *The Sunday Times book of brain teasers—Book 1.* London: Unwin. (Also published as *The Sunday Times book of brain teasers.* New York: St. Martin's Press, 1982.)

Bryant, V., & Postill, R. (1983). *The Sunday Times book of brain teasers—Book 2.* Englewood Cliffs, NJ: Prentice Hall.

Burkill, J. C., & Kundy, H. M. (1961). *Mathematical scholarship problems.* London: Cambridge University Press.

Butts, T. (1973). *Problem solving in mathematics.* Glenview, IL: Scott, Foresman.

CEMREL. (1975). *Elements of mathematics book, problem book* (2 vols.). St. Louis, MO: Author.

Clarke, B. R. (1994). *Puzzles for pleasure.* New York: Cambridge University Press.

Clarke, B. R., Gooch, R., Newing, A., & Singmaster, D. (1993). *The Daily Telegraph book of brain twisters No. 1.* London: Pan.

Conrad, S. R., & Flegler, D. (1992). *Math contests for high school* (2 vols.). Tenafly, NJ: Math League Press.

Conrad, S. R., & Flegler, D. (1992). *Math contests grades 7 and 8* (2 vols.). Tenafly, NJ: Math League Press.

Conrad, S. R., & Flegler, D. (1994). *Math contests grades 4, 5, and 6* (2 vols.). Tenafly, NJ: Math League Press.

Dorofeev, G., Potapov, M., & Rozov, N. (1973). *Elementary mathematics: Selected topics and problem solving.* Moscow: Mir Publishers.

Dorrie, H. (1965). *100 great problems of elementary mathematics.* New York: Dover.

Dowlen, N., Powers, S., & Florence, H. (1987). *College of Charleston mathematics contest books.* Palo Alto, CA: Dale Seymour Publications.

Dudney, H. E. (1926). *Modern puzzles.* New York: Pearson. (New edition, nd [1936])

Dudney, H. E. (1932). *Puzzles and curious problems.* London: Nelson.

Dudney, H. E. (1941). *A puzzle mine* (Ed. J. Travers). London: Nelson.

Dudney, H. E. (1958). *The Canterbury puzzles.* New York: Dover.

Dudney, H. E. (1967). *536 puzzles and curious problems* (Ed. Martin Gardner). New York: Scribner's.

Dudney, H. E. (1970). *Amusements in mathematics.* New York: Dover.

Dunn, A. (1964). *Mathematical bafflers.* New York: McGraw-Hill.

Dunn, A. F. (1983). *Second book of mathematical bafflers.* New York: Dover.

Dynkin, E. B., & Uspenskii, V. A. (1963). *Multicolor problems.* Heath.

Edwards, J. D., King, D. J., & O'Halloran, P. J. (1986). *All the best from the Australian Mathematics Competition.* Melbourne, Australia: Ruskin.

Emmet, E. R. (1976). *Mind tickling brain teasers.* Buchanan, NY: Emerson.

Emmet, E. R. (1976). *The Puffin book of brain teasers.* London: Puffin.

Emmet, E. R. (1977). *A diversity of puzzles.* New York: Barnes & Noble.

Emmet, E. R. (1977). *Puzzles for pleasure.* Buchanan, NY: Emerson.

Emmet, E. R. (1979). *The great detective puzzle book.* New York: Barnes & Noble.

Emmet, E. R. (1980). *The island of imperfection puzzle book.* New York: Barnes & Noble.

Emmet, E. R. (1984). *The Penguin book of brainteasers.* New York: Viking. (Compiled from Emmet's posthumous notes by David Hall & Alan Summers)

Emmet, E. R. (1993). *Brain puzzler's delight.* New York: Sterling.

Emmet, E. R., & Eperson, D. B. (1988). *Patterns in mathematics.* Oxford, UK: Blackwell.

Filipiak, A. S. (1942). *Mathematical puzzles.* New York: Bell.

Fisher, L., & Kennedy, B. (1984). *Brother Alfred Brousseau Problem Solving and Mathematics Competition, Introductory Division.* Palo Alto CA: Dale Seymour Publications.

Fisher, L., & Medigovich, W. (1984). *Brother Alfred Brousseau Problem Solving and Mathematics Competition, Senior Division.* Palo Alto, CA: Dale Seymour Publications.

Friedland, A. J. (1970). *Puzzles in math and logic.* New York: Dover.

Frohlichstein, J. (1962). *Mathematical fun, games and puzzles.* New York: Dover.

Fujimura, K. (1978). *The Tokyo puzzles* (Ed. by Martin Gardner). New York: Scribner's.

Gamow, G., & Stern, M. (1958). *Puzzle-math.* London: Macmillan.

Gardner, M. (1959). *Arrow book of brain teasers.* New York: Scholastic Press.

Gardner, M. (1959). *The Scientific American book of mathematical puzzles and diversions.* New York: Simon & Schuster. (Revised, with new afterword and references, as: *Hexaflexagons and other mathematical diversions.* Chicago: University of Chicago Press, 1988.)

Gardner, M. (1961). *The second Scientific American book of mathematical puzzles and diversions.* New York: Simon & Schuster.

Gardner, M. (1966). *Martin Gardner's new mathematical diversions from Scientific American.* New York: Simon & Schuster. (Reprinted in 1983 by the University of Chicago Press.)

Gardner, M. (1967). *The numerology of Dr. Matrix.* New York: Simon & Schuster.

Gardner, M. (1969). *Perplexing puzzles and tantalizing teasers.* New York: Simon & Schuster.

Gardner, M. (1969). *The unexpected hanging and other mathematical diversions.* New York: Simon & Schuster. (Revised ed. in 1991 by the University of Chicago Press.)

Gardner, M. (1971). *Martin Gardner's sixth book of mathematical games from Scientific American.* San Francisco: Freeman. (Reprinted in 1983 by the University of Chicago Press.)

Gardner, M. (1975). *Mathematical carnival.* New York: Knopf. (Revised ed. in 1989 by the Mathematical Association of America, Washington, DC.)

Gardner, M. (1976). *The incredible Dr. Matrix.* New York: Scribner's. (Contains all of *The Numerology of Dr. Matrix*)

Gardner, M. (1977). *Mathematical magic show.* New York: Knopf. (Revised ed. in 1990 by the Mathematical Association of America, Washington, DC.)

Gardner, M. (1977). *More perplexing puzzles and tantalizing teasers.* New York: Archway (Pocket Books).

Gardner, M. (1978). *Aha! Insight.* New York: Scientific American & Freeman.

Gardner, M. (1979). *Mathematical circus.* New York: Knopf. (Revised ed. in 1992 by the Mathematical Association of America, Washington, DC.)

Gardner, M. (1981). *Science fiction puzzle tales.* New York: C. N. Potter.

Gardner, M. (1982). *Aha! Gotcha.* New York: Freeman.

Gardner, M. (1983). *Wheels, life and other mathematical amusements.* New York: Freeman.

Gardner, M. (1985). *The magic numbers of Dr. Matrix.* Buffalo, NY: Prometheus, 1985. (Contains all of *The Incredible Dr. Matrix*)

Gardner, M. (1986). *Entertaining mathematical puzzles.* New York: Dover.

Gardner, M. (1986). *Knotted doughnuts and other mathematical entertainments.* New York: Freeman.

Gardner, M. (1986). *Puzzles from other worlds.* New York: Vintage (Random House).

Gardner, M. (1987). *Riddles of the sphinx.* Washington, DC: New Mathematical Library, Mathematical Association of America.

Gardner, M. (1988). *Time travel and other mathematical bewilderments.* New York: Freeman.

Gardner, M. (1989). *Penrose tiles to trapdoor ciphers.* New York: Freeman.

Gardner, M. (1992). *Fractal music, hypercards and more.* New York: Freeman.

Gardner, M. (1994). *My best mathematical and logical puzzles.* New York: Dover.

Garvin, A. D. (1975). *Discovery problems for better students.* Portland, ME: Weston Walch.

Gleason, A. M., Greenwood, R. E., & Kelly, L. M. (1980). *The William Lowell Putnam Mathematical Competitions. Problems and solutions: 1938-1964.* Washington, DC: Mathematical Association of America.

Gould, P. (1992). *Senior Challenge '85-'91: Mathematical education on Merseyside.* Liverpool, UK: University of Liverpool.

Gould, P., & Porteous, I. (1984). *Senior Challenge '80-'84: Mathematical education on Merseyside.* Liverpool, UK: University of Liverpool.

Graham, L. A. (1959). *Ingenious mathematical problems and methods.* New York: Dover.

Graham, L. A. (1968). *The surprise attack in mathematical problems.* New York: Dover.

Grietzer, S. L. (1978). *International Mathematical Olympiads, 1959-1977.* Washington, DC: Mathematical Association of America.

Haber, P. (1957). *Mathematical puzzles and pastimes.* Mount Vernon, NY: Peter Pauper.

Halmos, P. R. (1991). *Problems for mathematicians young and old* (Dolciani Mathematical Expositions #12). Washington, DC: Mathematical Association of America.

Higgins, A. M. (1971). *Geometry problems.* Portland, ME: J. Weston Walch.

Hill, T. J. (1974). *Mathematical Challenges II—Plus Six.* Washington, DC: National Council of Teachers of Mathematics.

Holton, D. (1988-1990). *Problem solving series.* Leicester, UK: Mathematical Association.

Honsberger, R. (1978). *Mathematical morsels.* Washington, DC: Mathematical Association of America.

Hunter, J. A. H. (1965). *Fun with figures.* New York: Dover.

Hunter, J. A. H. (1966). *More fun with figures.* New York: Dover.

Hunter, J. A. H. (1972). *Figures for fun.* London: Phoenix House.

Hunter, J. A. H. (1976). *Mathematical brain teasers* (Corrected and enlarged). New York: Dover.

Hunter, J. A. H. (1979). *Challenging mathematical teasers.* New York: Dover.

Hunter, J. A. H. (1983). *Entertaining mathematical teasers and how to solve them.* New York: Dover.

Kendall, P. M. H., & Thomas, G. M. (1962). *Mathematical puzzles for the connoisseur.* London: Griffin.

King, T. (c. 1927). *The best 100 puzzles solved and answered.* London: Foulsham.

Kinnaird, W. C. (Ed.). (1946). *Encyclopedia of puzzles and pastimes.* New York: Grosset & Dunlap.

Klamkin, M. S. (1986). *International Mathematical Olympiads, 1979-1985.* Washington, DC: Mathematical Association of America.

Kordemsky, B. A. *The Moscow puzzles* (Ed. Martin Gardner). New York: Scribner's.

Krechmer, V. A. (1974). *A problem book in algebra* (V. Shiffer, Trans.). Moscow: Mir Publishers.

Krulik, S., & Rudnick, J. A. (1980). *Problem solving: A handbook for teachers.* Boston: Allyn and Bacon.

Kutepov, A., & Rubanov, A. (1975). *Problems in geometry* (O. Meshkov, Trans.). Moscow: Mir Publisher.

Kutepov, A., & Rubanov, A. (1978). *Problem book: Algebra and elementary function* (L. Levant, Trans.). Moscow: Mir Publisher.

Larson, L. C. (1983). *Problem solving through problems.* New York: Springer-Verlag.

Lenchner, G. (1983). *Creative problem solving in school mathematics.* Boston: Houghton Mifflin.

Loyd, S. (1914). *Sam Loyd's cyclopedia of 5,000 puzzles, tricks and conundrums.* New York: Bigelow.

Loyd, S. (1927). *Sam Loyd's tricks and puzzles* (Vol. 1). New York: Experimenter Publishing.

Loyd, S. (1928). *Sam Loyd and his puzzles.* New York: Barse & Co.

Loyd, S. (1959). *Mathematical puzzles of Sam Loyd* (Vol. 1). New York: Dover.

Loyd, S. (1960). *Mathematical puzzles of Sam Loyd* (Vol. 2). New York: Dover.

Moser, W., & Barbeau, E. (1976). *The Canadian Mathematics Olympiads, 1969, 1975.* Montreal: Canadian Mathematical Congress.

Morris, I. (1969). *The Riverside puzzles.* New York: Walker & Co.

Morris, I. (1970). *The lonely monk and other puzzles.* Boston: Little, Brown.

Morris, I. (1972). *Foul play and other puzzles of all kinds.* New York: Vintage (Random House).

Moscovich, I. (1984). *Super-games.* London: Hutchinson.

Moscovich, I. (1991). *Fiendishly difficult math puzzles.* New York: Sterling.

Moscovich, I. (1991). *Fiendishly difficult visual perception puzzles.* New York: Sterling.

Moser, W. O. J., & Barbeau, E. J. (1978). *The first ten Canadian Mathematics Olympiads (1969-1978).* Montreal: Canadian Mathematical Society.

Mosteller, F. (1965). *Fifty challenging problems in probability.* New York: Dover.

Mott-Smith, G. (1954). *Mathematical puzzles for beginners and enthusiasts.* New York: Dover.

Newton, D. E. (1972). *One hundred quickies for math classes.* Portland ME: J. Weston Walch.

Phillips, H. (1932). *The week-end problems book.* London: Nonesuch.

Phillips, H. (1933). *The playtime omnibus.* London: Faber & Faber.

Phillips, H. (1934). *The sphinx problem book.* London: Faber.

Phillips, H. (1936). *Brush up your wits.* London: Dent.

Phillips, H. (1937). *Question time.* London: Dent.

Phillips, H. (1945). *Ask me another.* London: Ptarmigan.

Phillips, H. (1945). *Hubert Phillips's heptameron.* London: Eyre & Spottiswoode.

Phillips, H. (1945). *Something to think about.* London: Ptarmigan. (With additional Foreword, one problem omitted, and 11 problems added, London: Max Parrish, 1958.)

Phillips, H. (1947). *Playtime.* London: Ptarmigan.

Phillips, H. (1950). *The Hubert Phillips Annual 1951.* London: Hamish Hamilton.

Phillips, H. (1960). *Problems omnibus* (Vol. 1). London: Arco.

Phillips, H. (1961). *My best puzzles in logic and reasoning.* New York: Dover.

Phillips, H. (1961). *My best puzzles in mathematics.* New York: Dover.

Phillips, H. (1962). *Problems omnibus* (Vol. 2). London: Arco.

Phillips, H., Shovelton, S. T., & Marshal, G. S. (1961). *Caliban's problem book.* New York: Dover.

Polya, G., & Kilpatrick, J. (1974). *The Stanford mathematics book.* New York: Teachers College Press.

Posamentier, A. S. (1984). *Excursions in advanced Euclidean geometry.* Menlo Park, CA: Addison-Wesley.

Posamentier, A. S. (1996). *Getting ready for the SAT I—Math: Student guide with solutions for practice tests.* Thousand Oaks, CA: Corwin.

Posamentier, A. S., & Salkind, C. T. (1996). *Challenging problems in algebra* (Rev. ed.). New York: Dover.

Posamentier, A. S., & Salkind, C. T. (1996). *Challenging problems in geometry* (Rev. ed.). New York: Dover.

Posamentier, A. S., & Sheridan, G. (1984). *Math motivators: Pre-algebra, algebra, and geometry.* Menlo Park, CA: Addison-Wesley.

Posamentier, A. S., & Stepelman, J. (1995). *Teaching secondary school mathematics: Techniques and enrichment units* (4th ed.). Columbus, OH: Macmillan/Merrill.

Posamentier, A. S., & Wernick, W. (1988). *Advanced geometric constructions.* Palo Alto, CA: Dale Seymour Publications.

Ransom, W. R. (1955). *One hundred mathematical curiosities.* Portland, ME: J. Weston Walch.

Rapaport, E. (1963). *Hungarian problem book* (2 vols.). New York: Random House.

Reis, C. M., & Ditor, S. Z. (Eds.). (1988). *The Canadian Mathematics Olympiads (1979-1985).* Ottawa: Canadian Mathematical Society.

Ruderman, H. D. (1983). *NYSML-ARML Contests 1973-1982.* Norman, OK: Mu Alpha Theta.

Salkind, C. T. (1961). *The contest problem book.* New York: Random House.

Salkind, C. T. (1966). *The MAA problem book II.* New York: Random House.

Salkind, C. T., & Earl, J. M. (1973). *The MAA problem book III.* New York: Random House.

Saul, N. A., Kessler, G. W., Krilov, S., & Zimmerman, L. (1986). *The New York City Contest problem book.* Palo Alto, CA: Dale Seymour Publications.

Shklarsky, D. O., Chentzov, N. N., & Yaglom, I. M. (1979). *Selected problems and theorems in elementary mathematics* (V. M. Volosov and I. G. Volsova, Trans.). Moscow: Mir Publisher.

Shortz, W. (1991). *Will Shortz's best brain busters.* New York: Times Books (Random House).

Shortz, W. (1991). *Will Shortz's best brain twisters.* New York: Times Books (Random House).

Shortz, W. (1993). *Brain twisters from the First World Puzzle Championships.* New York: Times Books (Random House).

Sierpinski, W. (1964). *A selection of problems in the theory of numbers.* London: Pergamon/Macmillan.

Sierpinski, W. (1970). *250 problems in elementary number theory.* New York: American Elsevier.

Snape, C., & Scott, H. (1991). *How puzzling.* New York: Cambridge University Press.

Sole, T. (1988). *The ticket to heaven and other superior puzzles.* London: Penguin.

Sitomer, H. (1974). *The new mathlete problems book.* Valley Stream, NY: Nassau County Interscholastic Mathematics League.

Soifer, A. (1987). *Mathematics as problem solving.* Colorado Springs: Center for Excellence in Mathematics Education.

Steinhaus, H. (1963). *One hundred problems in elementary mathematics.* New York: Pergamon.

Straszewicz, S. (1965). *Mathematical problems and puzzles from the Polish Mathematical Olympiads* (J. Smsliska, Trans.). New York: Pergamon.

Trigg, C. W. (1967). *Mathematical quickies.* New York: McGraw-Hill.

Ulam, S. M. (1960). *Problems in modern mathematics.* New York: John Wiley.

Vout, C., & Gray, G. (1993). *Challenging puzzles.* New York: Cambridge University Press.

Wall, H. S. (1963). *Creative mathematics.* Austin: University of Texas Press.

Wells, D. (1982). *Can you solve these?* Norfolk, UK: Stradbroke.

Wells, D. G. (1979). *Recreations in logic.* New York: Dover.

Williams, W. T., & Savage, G. H. (1940). *The Penguin problems book.* London: Penguin.

Williams, W. T., & Savage, G. H. (c. 1940). *The Strand problems book.* London: Newnes.

Williams, W. T., & Savage, G. H. (1944). *The second Penguin problems book.* London: Penguin.

Williams, W. T., & Savage, G. H. (1946). *The third Penguin problems book.* London: Penguin.

Yaglom, A. M., & Yaglom, I. M. (1964). *Challenging mathematical problems with elementary solutions* (2 vols.). San Francisco: Holden-Day.

Readings on Problem Solving

Ackoff, R. L. (1978). *The art of problem solving.* New York: John Wiley.

Adams, J. L. (1974). *Conceptual blockbusting.* San Francisco: Freeman.

Adler, I. (1970). *Mathematics and mental growth.* London: Dobson.

Andre, T. (1986). Problem solving and education. In G. Phye & T. Andre (Eds.), *Cognitive classroom learning.* Orlando, FL: Academic Press.

Arnold, W. R. (1971). Students can pose and solve original problems. *The Mathematics Teacher, 64,* 325.

Averbach, B., & Chein, O. (1980). *Mathematics: Problem solving through recreational mathematics.* San Francisco: Freeman.

Bransford, J. D., & Stein, B. S. (1984). *The ideal problem solver.* New York: W. H. Freeman.

Brown, S. I., & Walter, M. I. (1983). *The art of problem posing.* Hillsdale, NJ: Lawrence Erlbaum.

Butts, T. (1985). In praise of trial and error. *The Mathematics Teacher, 78,* 167.

Charles, R., & Lester, F. (1982). *Teaching problem solving: What, why, and how.* Palo Alto, CA: Dale Seymour Publications.

Chipman, S., Segal, J., & Glaser, R. (1985). *Thinking and learning skills: Vol. 2. Research and open questions.* Hillsdale, NJ: Lawrence Erlbaum.

Cofman, J. (1990). *What to solve? Problems and suggestions for young mathematicians.* New York: Oxford University Press.

Costa, A. (1984, November). Mediating the metacognitive. *Educational Leadership,* pp. 57-62.

Curcio, F. (Ed.). (1987). *Teaching and learning: A problem solving focus.* Reston, VA: NCTM.

Davis, R., Jockusch, E., & McKnight, C. (1978). Cognitive processes in learning algebra. *Journal of Children's Mathematical Behavior, 2*(1).

Derry, S. J., & Murphy, D. A. (1986). Designing systems that train learning ability: From theory to practice. *Review of Educational Research, 56*(1), 1-39.

Emmet, E. R. (1981). *Learning to think.* Verplanck, NY: Emerson.

Fisher, R. B. (1981). *Brain games.* London: Fontana.

Fixx, J. F. (1978). *Solve it!* New York: Doubleday.

Frederiksen, N. (1984). Implications of cognitive theory for instruction on problem solving. *Review of Educational Research, 54*(3), 363-407.

Gardner, M. (1978). *Aha! Insight.* New York: Scientific American & Freeman.

Gardner, M. (1982). *Aha! Gotcha.* San Francisco: Freeman.

Gordon, W. J. J. (1961). *Synectics—The development of creative capacity.* New York: Harper & Row.

Hadamard, J. (1954). *The psychology of invention in the mathematical field.* New York: Dover.

Heiman, M., Narode, R., Slomianko, J., & Lochhead, J. (1987). *Thinking skills: Mathematics, teaching.* Washington, DC: National Education Association.

Honsberger, R. (1970). *Ingenuity in mathematics.* Washington, DC: Mathematical Association of America.

Honsberger, R. (1973). *Mathematical gems* (Vol. 1, Dolciani Mathematical Expositions #1). Washington, DC: Mathematical Association of America.

Honsberger, R. (1976). *Mathematical gems* (Vol. 2, Dolciani Mathematical Expositions #2). Washington, DC: Mathematical Association of America.

Honsberger, R. (1978). *Mathematical morsels* (Dolciani Mathematical Expositions #3). Washington, DC: Mathematical Association of America.

Honsberger, R. (1979). *Mathematical plums* (Dolciani Mathematical Expositions #4). Washington, DC: Mathematical Association of America.

Honsberger, R. (1985). *Mathematical gems III* (Dolciani Mathematical Expositions #9). Washington, DC: Mathematical Association of America.

Honsberger, R. (1991). *More mathematical morsels* (Dolciani Mathematical Expositions #10). Washington, DC: Mathematical Association of America.

Hough, J. S. (Ed.). *Problem solving: Newsletter* (Vol. 1-5). Philadelphia: Franklin Institute Press.

Hughes, B. (1975). *Thinking through problems.* Palo Alto, CA: Creative Publications.

Jensen, R. J. (1987). Stuck? Don't give up! Subgoal-generation strategies in rpoblem solving. *The Mathematics Teacher, 80,* 614.

Karmos, J., & Karmos, A. (1987). Strategies for active involvement in problem solving. In M. Heiman & J. Slomianko (Eds.), *Thinking skills instruction: Concepts and techniques* (pp. 99-110). National Educational Association.

Kluwe, R. (1987). Executive decisions and regulation of problem solving behavior. In F. Weinert & R. Kluwe (Eds.), *Metcognition, motivation and understanding.* Hillsdale, NJ: Lawrence Erlbaum.

Krulik, S. (Ed.). (1980). *Problem solving in school mathematics, 1980 yearbook.* Reston, VA: National Council of Teachers of Mathematics.

Krulik, S., & Rudnick, J. (1987). *Problem solving: A handbook for teachers* (2nd ed.). Boston: Allyn and Bacon.

Krulik, S., & Rudnick, J. (1989). *Problem solving: A handbook for senior high school teachers.* Boston: Allyn and Bacon.

Krulik, S., & Rudnick, J. (1993). *Reasoning and problem solving: A handbook for elementary school teachers.* Boston: Allyn and Bacon.

Krulik, S., & Rudnick, J. (1995). *The new sourcebook for teaching reasoning and problem solving in elementary schools.* Boston: Allyn and Bacon.

Krulik, S., & Rudnick, J. (1996). *The new sourcebook for teaching reasoning and problem solving in secondary schools.* Boston: Allyn and Bacon.

Mason, J. (1978). *Learning and doing mathematics.* Milton Keynes, UK: Open University Press.

Mason, J., with Burton, L., & Stacey, K. (1985). *Thinking mathematically.* Reading, MA: Addison-Wesley.

Mayer, R. (1986). Mathematics. In R. Dillon & R. Sternberg (Eds.), *Cognition and instruction.* Orlando, FL: Academic Press.

McKim, R. H. (1980). *Thinking visually: A strategy manual for problem solving.* Palo Alto, CA: Dale Seymour Publications.

Moses, S. (1974). *The art of problem-solving.* London: Transworld.

Mottershead, L. (1978). *Sources of mathematical discovery.* Oxford, UK: Blackwell.

Mottershead, L. (1985). *Investigations in mathematics.* Oxford, UK: Blackwell.

Nickerson, R. (1981, October). Thoughts on teaching thinking. *Educational Leadership,* pp. 21-24.

Nickerson, R., Perkins, D., & Smith, E. (1985). *The teaching of thinking.* Hillsdale, NJ: Lawrence Erlbaum.

Polya, G. (1945). *How to solve it.* Princeton, NJ: Princeton University Press.

Polya, G. (1954). *Introduction and analogy in mathematics.* Princeton, NJ: Princeton University Press.

Polya, G. (1954). *Patterns of plausible inference.* Princeton, NJ: Princeton University Press.

Polya, G. (1962). *Mathematical discovery* (2 vols.). New York: John Wiley. (Combined ed. with foreword by Peter Hilton, bibliography extended by Gerald Alexanderson, and index extended by Jean Pedersen, New York: John Wiley, 1981.)

Posamentier, A. S., & Schulz, W. (Eds.). (1996). *The art of problem solving: A resource for the mathematics teacher.* Thousand Oaks, CA: Corwin.

Reeves, C. A. (1987). *Problem solving techniques helpful in mathematics and science.* Reston, VA: National Council of Teachers of Mathematics.

Schoenfeld, A. H. (1983). *Problem solving in the mathematics curriculum.* Washington, DC: Mathematical Association of America.

Schoenfeld, A. H. (1985). *Mathematical problem solving.* Orlando, FL: Academic Press.

Segal, J., Chipman, S., & Glaser, R. (Eds.). (1985). *Thinking and learning skills. Vol. 1: Relating instruction to research.* Hillsdale, NJ: Lawrence Erlbaum.

Silver, E. A. (Ed.). (1985). *Teaching and learning mathematical problem solving.* Hillsdale, NJ: Lawrence Erlbaum.

Simon, M. A. (1986). The teacher's role in increasing student understanding of mathematics. *Educational Leadership, 43*(7), 40-43.

Skemp, R. R. (1971). *The psychology of learning mathematics.* Baltimore, MD: Penguin.

Smullyan, R. (1978). *What is the name of this book?* Englewood Cliffs, NJ: Prentice Hall.

Soifer, A. (1987). *Mathematics as problem solving.* Colorado Springs: Center for Excellence in Mathematics Education.

Topoly W. (1965). An introduction to solving problems. *The Mathematics Teacher, 58,* 48.

Troutman, A., & Lichtenberg, B. P. (1974). Problem solving in the general mathematics classroom. *The Mathematics Teacher, 67,* 590.

Walter, M. I., & Brown, S. I. (1977). Problem posing and problem solving. *The Mathematics Teacher, 70,* 4.

Whirl, R. J. (1973). Problem solving—Solution or technique? *The Mathematics Teacher, 66,* 551.

Winckelgren, W. A. (1974). *How to solve problems.* San Francisco: W. H. Freeman.

About the Editors

Alfred S. Posamentier is Professor of Mathematics Education and Associate Dean of the School of Education of the City College of the City University of New York. He is the author and coauthor of numerous mathematics books for teachers and secondary school students. As a very sought-after guest lecturer, he favors topics regarding aspects of mathematics problems solving and the introduction of uncommon topics into the secondary school realm for the purpose of enriching the mathematics experience of those students. After completing his A.B. degree in mathematics at Hunter College of the City University of New York, he took a position as a teacher of mathematics at Theodore Roosevelt High School in the Bronx (New York). During his 6 years in that position, he focused his attention on improving the students' problem-solving skills. Immediately upon joining the faculty of the City College (after having received his Master's degree there), he began to develop in-service courses for secondary school mathematics teachers on the topic of mathematics problem solving. These courses were extremely popular, drawing teachers from outside the usual attendance area. Dr. Posamentier received his Ph.D. from Fordham University (New York) in mathematics education and since has extended his reputation to Europe. He has been a visiting professor at several British, German, and Austrian universities. Dr. Posamentier is often cited for his outstanding teaching. He was recently named Educator of the Year by the City College Alumni Association.

Wolfgang Schulz is a Professor of Mathematics Education in the Department of Mathematics at the Humboldt University of Berlin. He began his career after majoring in mathematics and geography at the Humboldt University, as a mathematics teacher in Berlin at the junior and senior high school level. During his time as a teacher of mathematics, his interest in problem solving developed as did his interest in mathematics education on a more theoretical level. This interest led him to work closely with students preparing for mathematics olympiads and other mathematics contests. Student success followed directly! For many years now, Professor Schulz has been concentrating his efforts and interests in the theoretical and practical aspects of the instruction of mathematics as well as the preparation of mathematics teachers. Professor Schulz is an author/editor of numerous German mathematics textbooks for secondary and upper-elementary school students. He also is coauthor of the textbook *Gruppoide und Funktionalgleichungen* designed to train mathematics teachers. As an author of numerous journal articles, he frequently lectures at conferences throughout Europe. Throughout his career, Dr. Schulz has worked with highly motivated students in the area of problem solving, and he has had numerous enriching experiences in which students stumble onto uniquely clever solutions to relatively common problems.

About the Contributors

Ethan Akin received his B.S. degree in 1965 from the City College of New York, where he is now Professor of Mathematics. His many research papers and expository articles include such titles as "Recurrence of the Unfit," "Bulgarian Solitaire," "Ovulation Regulation: The Right Number or Nothing," and "The Spiteful Computer," as well as monographs: *The Metric Theory of Banach Manifolds* and *The Geometry of Population Genetics*. His comprehensive survey, *The General Topology of Dynamical Systems* was published by the American Mathematical Society in 1993. Professor Akin has helped design an alternative problem-solving calculus course and taught one of its initial two sections. The psychology of teaching and learning remains one of his strong interests.

Steven R. Conrad has taught since 1980 at Roslyn High School, Roslyn Heights, New York. Prior to that, he taught at Benjamin Cardozo High School in Bayside and Francis Lewis High School, in Flushing—both in New York City. He received a B.S. from Queen's College, an M.S. from Yeshiva University, and a certificate in School Administration from St. John's University. He has done additional graduate work at the University of San Francisco and Fordham University. A speaker at numerous international, national and regional conferences, Mr. Conrad has also been a contributor to the problem sections of many mathematics journals. He has served as Editor of *Summation* and Associate Editor of *The New York State Mathematics' Teachers' Journal* and *The Mathematics Student Journal*. He has had articles published in *Mathematics Magazine* and *The Mathematics Teacher*. Mr. Conrad has coauthored six books of mathematics contests.

Ira Ewen has taught mathematics at the secondary and college levels ever since he earned his A.B. and MA degrees in pure mathematics in 1952 and 1953 from Harvard University. He has served as a teacher of mathematics, an assistant principal supervising a high school mathematics department, and a high school principal. He has been a teaching fellow in mathematics at Harvard University, a visiting assistant professor at the University of Illinois, and an adjunct assistant professor at the City College of New York. During the 1980s he served as Director of Thinking Skills for the Division of Curriculum and Instruction and designed the New York City model of the Reasoning Skills Project in Mathematics and English under a multicity initiative of the college board. He developed many strategies for improving student ability to reason effectively and modeled the use of games as a way of stimulating higher-order thinking in students at all levels of ability, ranging from those with learning disabilities to the academically gifted and talented.

Lenni I. Haapasalo received his Ph.D. in Mathematics at the University of Jyvaskyla in Finland. After having taught mathematics for 9 years in the Department of Mathematics, he moved into the Department of Teacher Education, where he serves as Docent of Mathematics Didactics. He has written several books and more than 60 articles of mathematics education in Finnish, English, German, or Swedish. He has also written three modern textbooks to be used in comprehensive and vocational schools and about 20 programs for computer assisted learning. His research in modern constructivist learning has led him to develop models for the systematic planning of learning environments, and this project has actually grown into a development program for Finnish mathematics education.

Hope J. Hartman is Associate Professor, head of Social and Psychological Foundations of Education, and Director of the Tutoring and Cooperative Learning Program at the City College of the City University of New York. Hartman's areas of specialization are improving thinking skills, improving attitudes that foster thinking and learning, tutoring, and cooperative learning. She received a graduate fellowship and earned her Ph.D. in Cognitive Psychology at Rutgers—Newark in New Jersey. In 1984, Hartman moved to Montclair State College to become Assistant Director of the Thinking Skills Program "Project THISTLE" and Assistant Professor in the School of Education. She began focusing her improving thinking approach to improving students' academic performance on tutoring and cooperative learning in 1986 when she came to The City College of the City University of New York.

Hans Humenberger is a high school teacher in Vienna and holds a master's degree in Mathematics from the University of Vienna. His interest in the use of applications to teach mathematics brought him to work as an assistant in the Institute of Mathematics at the University of Vienna under the auspices of Professor Reichel. There he was involved in the training of student teachers and pursued the study of the relationship between language and mathematics. His main interest is problem solving in mathematics. In 1992, he was awarded a doctorate from the Department of Mathematics at the University of Vienna.

Hans K. Kaiser is a Professor of Mathematics at the University of Technology of Vienna (Austria). Apart from mathematics he has always been interested in history. To support his studies, he worked as a part-time teacher. In his professional career, he has tried to combine all these interests. His fields of research are universal algebra, applications of algebra in computer science, and the history and didactics of mathematics. He has published a number of research articles in these areas and is coauthor of several books (*Algebra für Informatiker* [*Algebra for Computer Scientists*] for university students and *Geschichte der Mathematik im Schulunterricht* [*History of Mathematics for the Schools*], which is intended for teachers of secondary schools). At the University of Technology of Vienna, he teaches mathematics to students who want to become engineers or teachers in mathematics and computer science.

Marion Kauke is a member of the Pedagogical Psychology faculty in the Department of Education at Humboldt University of Berlin. At the Friedrich Schiller University in Jena, Germany, she majored in Social Psychology with a specialty in research on the theoretical bases for game experiments/strategies, graduated with honors in 1972, and earned her doctorate in 1975. At the age of 33, she became the youngest female university faculty member in the former GDR. She instructed inventors and patent engineers in the area of creativity and taught soccer players as well as the GDR national chess team, where she served as the official psychologist. At the Educational Institution for Creativity Studies in Leipzig and at the College of Fine Arts in Berlin, Dr. Kauke gave courses in game theory. At the same time as her well received book *Game Intelligence: Playful Learning—Teaching Play?* was published in 1992, she directed a project with support from the German Research Association titled "Cooperation, Competition and Conflictual Relations Among School Children of the Same Age" as part of a worldwide cultural comparison.

Karl Kiesswetter was born in Sudetendland and has studied mathematics at the University of Koln. He was a high school teacher and then held positions at the Universities of Munster, Bielefeld, and Hamburg. He has been interested in mathematics education and fostering creativity among gifted school children ("Hamburg model").

Stephen Krulik is Professor and Coordinator of Mathematics Education at Temple University in Philadelphia, where he is responsible for the preparation of mathematics teachers of Grades K-12 and the in-service training of mathematics teachers at the graduate level. He taught mathematics in the New York City public schools for 15 years before coming to Temple. Nationally, he served as a mender of the committee responsible for preparing the *Professional STANDARDS for Teaching Mathematics* of the National Council of Teachers of Mathematics and was the editor of their 1980 yearbook, *Problem Solving in School Mathematics*. Regionally, he was President of the Association of Mathematics Teachers of New Jersey and served on the writing team that produced their 1993 publication, *The New Jersey Calculator Handbook*. Together with Jesse Rudnick, he has just coauthored *The New Sourcebook for Teaching Reasoning and Problem Solving in Elementary School*.

Wei Lee currently teaches mathematics at the City College of New York. He received his B.A. from the Shanghai University of Science and Technology. Upon graduation, he worked at the Computer Research Institute, when, after 4 years of problem-solving experiences of a somewhat different nature, he emigrated to the United States, where he is currently pursuing his doctoral studies while continuously looking for novel mathematical problem solving experiences. His contributions to this book exhibit his interest in using combinatorics and graph theory to solve routine and nonroutine mathematical problems.

Evan M. Maletsky joined the Department of Mathematics at Montclair State University in 1957, earning his bachelor's and master's degrees from there and his doctorate from New York University. His long tenure at Montclair State includes extensive experience in teaching a wide variety of mathematics and mathematics education courses at both the undergraduate and graduate levels and junior high school mathematics in the demonstration school that used to be on campus. He has served as consultant for and has conducted many in-service workshops in school districts in his home state of New Jersey and throughout the country. In addition, he is a frequent speaker at professional meetings. Maletsky's special interests include geometry and its relationship to manipulative, visualization, and problem-solving experiences and activities in the teaching and learning of mathematics. His numerous publications include a professional book on the teaching of mathematics and several mathematics textbook series at the elementary, middle, junior, and senior high levels. Current interests include the study of fractals and their connection to the school curriculum, the coauthoring of a series of publications titled *Fractals for the Classroom: Strategic Activities*, and editorial advising and writing on an upcoming middle-school mathematics textbook series.

Stephen E. Moresh is Supervisor of the Mathematics and Computer Science Department at Seward Park High School in New York, New York. He has taught on both the high school and university levels, having adjuncted at several colleges in the New York Metropolitan area. In the early 1980s, he served as the director of the Mathematics-Science Recertification Program at Teachers College, Columbia University and now teaches graduate mathematics courses in the School of Education at City College of New York, where he became interested in problem solving in combinatorics. Dr. Moresh has coauthored textbooks in computer mathematics and algebra (using an integrated approach) and several curriculum guides for the New York City Board of Education and has written several articles for local publications. He earned his Baccalaureate and Master's degrees at Brooklyn College (CUNY) and a second Master's degree and doctorate at Columbia University, New York. He was honored by being a state award winner in 1983 of the Presidential Awards for Excellence in Science and Mathematics Teaching and has also served as a reader for the Advanced Placement Calculus Examinations.

Fred Paul was with the Bureau of Mathematics of the New York State Education Department for 27 years until his retirement in October 1992. For the last 17 of these years he served as Bureau Chief, and in that capacity he directed all aspects of mathematics curricula and assessment for Grades K-12. He was instrumental in the development and implementation of a statewide integrated and sequential secondary mathematics program and in the writing and distribution of numerous teacher-support teacher-support publications. Among these publications were several on creative problem solving. Paul previously taught at the junior high, senior high, and college levels, and conducted in-service courses for elementary teachers. He received B.A. and MA degrees and pursued further graduate work at the New York State College for Teachers, now the University at Albany. Since his retirement, he has been active in mathematics education as a consultant for several school districts, as a speaker at conferences and meetings, and as an in-service course instructor.

Hans Christian Reichel is Professor of Mathematics and Mathematics Education at the Institute of Mathematics at the University of Vienna. He is a recognized expert in topology and has given many exciting lectures on the philosophy of mathematics. He also has a keen interest in the teaching of mathematics, especially to students in the 10- to 18-year-old age bracket. Using his interest in application-oriented mathematics, he has coauthored 12 volumes of schoolbooks, which are currently being used throughout Austria. Dr. Reichel received a master's degree in physics and mathematics from the University of Vienna in 1968 and was awarded a doctorate in mathematics from that institution in 1969.

Brigitte Rollett is Professor of Psychology and Head of the Department of Developmental and Educational Psychology at the University of Vienna/Austria. After finishing her doctoral studies at the University of Graz, she held positions as a full professor at the Teacher's Training College at Osnabruck, the University of Kassel, and the University of Bochum. She is President of the Austrian Psychological Association and National Correspondent of EARLI (European Association of Research in Learning and Instruction), the European twin organization of AERA, has published over 120 articles and chapters in handbooks, and is the author/editor of 11 books. Her main research interests lie in the field of school and life achievement and its developmental and experiential causes.

Jesse A. Rudnick is Professor Emeritus at Temple University in Philadelphia, where he was Professor and Coordinator of Mathematics Education. Prior to coming to Temple, he was a teacher of mathematics in the secondary schools of Philadelphia for 15 years and Associate Professor of Mathematics at Glassboro State College (now Rowan College) in New Jersey. Nationally, he has served as a director of the National Council of Teachers of Mathematics, as well as chairman of their Conferences and Conventions Committee and Mathematics Education Trust. He has been a regional director of the National Council of Supervisors of Mathematics, and locally, was President of the Association of Mathematics Teachers of Philadelphia and Vicinity. His awards include the "Outstanding Contribution to Mathematics Education in Pennsylvania" award and the "Outstanding Contributions to Mathematics Education" award.

Mario Salvadori, a mathematical physicist and an engineer, is James Renwick Professor Emeritus of Civil Engineering and Professor Emeritus of Architecture, Columbia University, where he served for 50 years on three of its faculties. He has served as Professor of Architecture at Princeton University and has lectured from Japan to Brazil and from Taiwan to most of the European countries. His scientific output consists of over 150 technical papers in American magazines and magazines of other countries and 18 books on engineering mathematics, structural theory, architecture and aesthetics. His engineering forensic activity covers 50 years in the United States courts, analyzing and explaining the cause of structural collapses.

David Breyer Singmaster is Professor of Mathematics at the School of Computing, Information Systems and Mathematics, at South Bank University in London. His main fields of interest include recreational mathematics and its history; popularization of mathematics (conveying the excitement and enjoyment of mathematics to students, teachers, and the public); combinatorics; elementary number theory; computing and its history; history of science and technology; and polyhedra and convex sets. In 1976, he became interested in encouraging the use of mathematical films and compiled a *List of 16 mm Films on Mathematical Subjects,* which went through five editions in 1979-1983. From 1978 to about 1984, he was the leading expositor of Rubik's Cube and its successors, particularly the mathematics involved, and wrote the first book on the Cube, which went into five editions.

Roland Stowasser is Professor at the Department of Mathematics of the Technische Universität in Berlin. He taught for several years at the Carl-Duisberg- Gymnasium. He later served as Director of the Studienseminar in Wuppertal (Germany) and Associate Professor at the Universität Bielefeld (Institut fur Didaktik der Mathematik). He has coauthored secondary school textbooks, published papers in a variety of educational journals, and served as chief editor of the German journal *Mathematiklehrer* (The Mathematics Teacher). Professor Stowasser's primary research interests are the history of mathematics as related to mathematics education, curriculum development in the light of microcomputer technology, and problem solving and artificial intelligence. Since 1991 he has been engaged in the European joint] project "New Approaches to the Teaching of Engineering Mathematics," giving low-level mathematical modeling and the art of visual problem posing his special attention.

Jan Troják was born in Prague, Bohemia, where he also received his university education. In 1961, he graduated from the Faculty of Science of the Pedagogical University with a diploma of a teacher of mathematics and descriptive geometry. Since that time he worked and studied at the Faculty of Mathematics and Physics of the Charles University in Prague. Between 1968-1971 and 1973-1975, he worked as a lecturer at the College of Science of the University of Basrah and at the University of Baghdad in Iraq, giving lectures on projective, non-Euclidean and differential geometries, graph theory, calculus and analytical geometry, and linear algebra. Dr. Troják has cowritten a textbook (published in three volumes) on analytical geometry of the Euclidean, affine spaces, and affine and congruence transformations.

Sabine Ziller received a B.S. in Mathematics Statistics and Probability at the Humboldt University, Berlin in 1982. She has also received university degrees in Educational Psychology (1989), Piano (at the Music Conservatory, 1976), and Music Education (1991). She worked as a research assistant at the Academy of Science and directed the computer division at an East Berlin hospital. Throughout this time, she also conducted a project on singing and playing for elementary school students. There she employed her mathematics and musical talents to the creative advantage of young children. Currently, she is working on a research project at the German Research Society titled "Cooperation, Competition and Conflict Relationships for Similar Aged Students in East Berlin." This project is part of a worldwide study on a similar theme.

CORWIN
PRESS

The Corwin Press logo—a raven striding across an open book—represents the happy union of courage and learning. We are a professional-level publisher of books and journals for K-12 educators, and we are committed to creating and providing resources that embody these qualities. Corwin's motto is "Success for All Learners."

The Art of Problem Solving

A Resource for the Mathematics Teacher

Editor:
Alfred S. Posamentier
The City College of the City University of New York

Associate Editor:
Wolfgang Schulz
Humboldt University, Berlin

As a mathematics teacher, you *know* how hard it can be for your students to understand and solve math problems. The old ways of problem solving don't always work. Even the most innovative teachers need some fresh ideas to make mathematics something students comprehend and enjoy.

Problem solving is the main theme of this new, idea-filled handbook. The chapter authors look at the subject in a completely new light, and the result is an enticing, entertaining, and *useful* resource.

The editors of this guidebook present a host of interesting ideas that range from practical to theoretical, from common to glitzy, that you can adapt for use in your classroom. With margin notes identifying subject matter and strategy type, this book is as easy to read as it is valuable.

The authors discuss problem-solving techniques such as:

- Solving a simpler analogous problem
- Using visual representation
- Working backwards
- Using a computer
- Considering extreme cases

Mathematics professionals from all over the world bring you their personal favorite strategies for problem solving. The chapters are independent of each other, and the editors encourage you to use the chapters in whatever order suits your needs.

If you have ever despaired of making mathematics into a favorite subject for yourself and your class, this is the book you need. These passionate professionals give you new insight into the art of problem solving.

Cover design by Marcia R. Finlayson

ISBN 0-8039-6362-9
9 780803 963627
90000>

CORWIN PRESS, INC.
A Sage Publications Company
2455 Teller Road
Thousand Oaks, CA 91320-2218
e-mail: order@corwin.sagepub.com
Call: 805-499-9774 Fax: 805-499-0871

BEST COPY AVAILABLE